



Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
|||||  
Db 7 ASREAKKQVEKALE 20

RESULT 2  
AAR97390  
ID AAR97390 standard; peptide; 20 AA.  
XX  
AC AAR97390;  
XX  
DT 27-AUG-2003 (revised)  
DT 02-DEC-1996 (first entry)  
XX  
DE Streptococcal M protein peptide p145, for use in chimaeric peptide.  
XX  
KM Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
KM conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KM diagnosis.  
XX  
OS Streptococcus sp.  
XX  
PN MO9611944-A1.  
XX  
PD 25-APR-1996.  
XX  
PR 16-OCT-1995; 95WO-AU000681.  
XX  
PR 14-OCT-1994; 94AU-00008851.  
XX  
PA (COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALT INST MEDICAL RES WALTER & ELITZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX  
DR WPI; 1996-221939/22.  
XX  
PT New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Claim 4; Page 80; 99pp; English.  
XX  
CC The present peptide is the Streptococcal M protein peptide p145  
CC (Pukaarn et al, J. Immunol. 149: 2729-2735 (1992)), used in the  
CC construction of a novel chimaeric peptide (CP). The CP comprises a B-cell  
CC conformational epitope from within the present peptide, inserted into a  
CC 2nd peptide, pref. an alpha-helical coil based on the GCN4 leucine zipper  
CC peptide (AAR97395). The 2nd peptide has a similar conformation, enabling  
CC the epitope to be presented in an immunologically active conformation.  
CC The CP can be used in a novel detection/mapping process, e.g. to  
CC determine the min. epitope required to induce opsonic antibodies (Ab),  
CC and in vaccines against gp. A Streptococci. Ab raised against the CP can  
CC be used for immunotherapy and diagnosis, while the CP can be used  
CC diagnostically to detect Ab. (Updated on 27-AUG-2003 to correct OS  
CC field.)  
XX  
SQ Sequence 20 AA;  
XX  
Query Match 100.0%; Score 64; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
|||||  
Db 7 ASREAKKQVEKALE 20

Db 7 ASREAKKQVEKALE 20

RESULT 3  
AAM69281  
ID AAM69281 standard; peptide; 20 AA.  
XX  
AC AAM69281;  
XX  
DT 29-OCT-1998 (first entry)  
XX  
DE Streptococcus group A protein fragment.  
XX  
KM Acryloylated peptide polymer; immune response; peptide epitope;  
KM synthetic vaccine; enzymatically cleavable site.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Msc-difference 1 /note= "linked to acryloylated peptide polymer"  
FT  
XX  
PN WO9834968-A1.  
XX  
PD 13-AUG-1998.  
XX  
PE 10-FEB-1998; 98WO-AU000076.  
XX  
PR 11-FEB-1997; 97AU-00005071.  
PR 03-OCT-1997; 97CA-02217321.  
XX  
PA (COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALT INST MEDICAL RES WALTER & ELITZA.  
PA (CSLC-) CSL LTD.  
XX  
PI Jackson DC, Obrien-Simpson NM, Brown LE, Zeng W, Ede NJ;  
PI Brandt ER, Good MF;  
XX  
DR WPI; 1998-447177/38.  
XX  
PT Acryloylated peptide polymers - useful for synthetic vaccine technology,  
PT for raising an immune response to peptide epitope and as diagnostic tool.  
XX  
PS Example 1; Page 20; 77pp; English.  
XX  
CC This sequence represents a fragment of a Streptococcus group A protein.  
CC This sequence was used to test the acryloylated peptide polymer of the  
CC invention. The peptide polymers are used to raise an immune response to a  
CC peptide epitope (such as this sequence), and also as diagnostic tools.  
CC Polymers (molecular wt. >600 kDa.) can be prepared with virtually any  
CC number of the same or different epitopes by a method that allows  
CC purification of the individual determinants, avoids errors inherent in  
CC long sequential syntheses in which protected peptide fragments are not  
CC used, thus avoiding solubility and purification problems. Multiple copies  
CC of many different peptide epitopes may be incorporated into a single  
CC polymeric structure to allow utilisation of the range of T cell epitopes  
CC required for outbred populations in conjunction with epitopes  
CC representing different pathogenic serotypes, thus making them a  
CC significant advance in synthetic vaccine technology  
XX  
SQ Sequence 20 AA;  
XX  
Query Match 100.0%; Score 64; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
|||||  
Db 7 ASREAKKQVEKALE 20

RESULT 4  
ADK00565  
ID ADK00565 standard; peptide; 29 AA.  
XX  
XX  
AC ADK00565;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Immunogenic lipopeptide of the invention #101.  
XX  
KW T helper cell epitope; B cell epitope; Antibacterial; Antitumor;  
KW Antinfertility; Vaccine; antibody.  
XX  
OS Synthetic.  
XX  
PN WO2004014956-A1.  
XX  
PD 19-FEB-2004.  
XX  
PF 12-AUG-2003; 2003WO-AU001018.  
XX  
PR 12-AUG-2002; 2002US-0402838P.  
XX  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
PI Jackson D, Zeng W;  
XX  
DR WPI; 2004-238735/22.  
XX  
PT Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.  
XX  
XX  
PS Claim 29; SEQ ID NO 101; 194pp; English.  
XX  
CC The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC kegr; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.  
XX  
XX  
SQ Sequence 29 AA;  
XX  
Query Match 100.0%; Score 64; DB 8; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ASREAKKQVERALE 14  
DB 9 ASREAKKQVERALE 22

RESULT 5  
ADK00571  
ID ADK00571 standard; peptide; 45 AA.  
XX  
XX  
AC ADK00571;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Immunogenic lipopeptide of the invention #107.  
XX  
KW T helper cell epitope; B cell epitope; Antibacterial; Antitumor;

KW Antinfertility; Vaccine; antibody.  
XX  
XX  
OS Synthetic.  
XX  
XX  
PN WO2004014956-A1.  
XX  
PD 19-FEB-2004.  
XX  
PF 12-AUG-2003; 2003WO-AU001018.  
XX  
PR 12-AUG-2002; 2002US-0402838P.  
XX  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
PI Jackson D, Zeng W;  
XX  
DR WPI; 2004-238735/22.  
XX  
XX  
PT Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.  
XX  
XX  
PS Claim 39; SEQ ID NO 107; 194pp; English.  
XX  
XX  
CC The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC kegr; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.  
XX  
XX  
SQ Sequence 45 AA;  
XX  
Query Match 100.0%; Score 64; DB 8; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ASREAKKQVERALE 14  
DB 25 ASREAKKQVERALE 38

RESULT 6  
ADK00569  
ID ADK00569 standard; peptide; 46 AA.  
XX  
XX  
AC ADK00569;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Immunogenic lipopeptide of the invention #105.  
XX  
KW T helper cell epitope; B cell epitope; Antibacterial; Antitumor;  
KW Antinfertility; Vaccine; antibody.  
XX  
OS Synthetic.  
XX  
PN WO2004014956-A1.  
XX  
PD 19-FEB-2004.  
XX  
PF 12-AUG-2003; 2003WO-AU001018.  
XX  
PR 12-AUG-2002; 2002US-0402838P.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA Jackson D, Zeng W;  
PI WPI; 2004-238735/22.  
DR  
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.  
XX  
PS Claim 39; SEQ ID NO 105; 194pp; English.  
XX  
CC The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC legr; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.  
XX  
SQ Sequence 46 AA;  
Query Match 100.0%; Score 64; DB 8; Length 46;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ASREAKKQVEKALE 14  
DB 26 ASREAKKQVEKALE 39  
RESULT 7  
ADK00572  
ID ADK00572 standard; peptide; 46 AA.  
XX  
AC ADK00572;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Immunogenic lipopeptide of the invention #108.  
XX  
KM T helper cell epitope; B cell epitope; Antibacterial; Antiulcer;  
KM Antifertility; Vaccine; antibody.  
XX  
OS Synthetic.  
XX WO2004014956-A1.  
XX  
PN 19-FEB-2004.  
XX  
PF 12-AUG-2003; 2003WO-AU001018.  
XX  
PR 12-AUG-2002; 2002US-0402838P.  
XX  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
PI Jackson D, Zeng W;  
XX  
DR WPI; 2004-238735/22.  
XX  
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.

XX Claim 39; SEQ ID NO 108; 194pp; English.  
PS  
XX  
CC The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC legr; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.  
XX  
SQ Sequence 46 AA;  
Query Match 100.0%; Score 64; DB 8; Length 46;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ASREAKKQVEKALE 14  
DB 26 ASREAKKQVEKALE 39  
RESULT 8  
ADK00570  
ID ADK00570 standard; peptide; 47 AA.  
XX  
AC ADK00570;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Immunogenic lipopeptide of the invention #106.  
XX  
KM T helper cell epitope; B cell epitope; Antibacterial; Antiulcer;  
KM Antifertility; Vaccine; antibody.  
XX  
OS Synthetic.  
XX WO2004014956-A1.  
XX  
PN 19-FEB-2004.  
XX  
PD 12-AUG-2003; 2003WO-AU001018.  
XX  
PF 12-AUG-2002; 2002US-0402838P.  
XX  
PR (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
PA Jackson D, Zeng W;  
XX  
PI WPI; 2004-238735/22.  
XX  
DR  
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.  
XX  
PS Claim 39; SEQ ID NO 106; 194pp; English.  
XX  
CC The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC legr; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for



CC antibody production, synthetic vaccine production, diagnostic method  
 CC employing antibodies and antibody ligands and immunotherapy for  
 CC veterinary and human medicine. The method efficiently elicits the  
 CC production of antibody against antigenic B cell epitope. The present  
 CC sequence represents a novel immunogenic lipopeptide comprising T helper  
 CC and B cell epitopes.

XX  
 SQ Sequence 47 AA;

Query Match 100.0%; Score 64; DB 8; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKALE 14  
 |||||  
 Db 27 ASREAKQVEKALE 40

RESULT 9

AAR10221  
 ID AAR10221 standard; protein; 234 AA.

XX  
 AC AAR10221;

XX 25-MAR-2003 (revised)

DT 26-MAR-1991 (first entry)

XX Streptococcal M6' protein.

XX Streptococcal M protein; M'6 protein; vaccinia virus; fowlpox virus;

XX poxviridae vaccine; streptococcal pharyngitis.

OS Streptococcus sp.

PN W09015872-A.

XX 27-DEC-1990.

PF 21-JUN-1989; 89US-00369118.

PR 21-JUN-1989; 89US-00369118.

PR 19-JUN-1990; 90US-00540586.

XX (UYRQ ) UNIV ROCKEFELLER.  
 PA (UYOR-) UNIV OREGON STATE.

XX Flischetti VA, Kruby DE;

XX WPI; 1991-022236/03.

DR N-PSDB; AAQ10244.

PT New recombinant streptococcal M protein DNA and viral vector - for  
 PT production of poxviridae vaccines in treatment of vaccinia, fowlpox etc.

XX Disclosure; Fig 5; 41pp; English.

CC This M'6 protein corresponds to the conserved exposed polypeptide region  
 CC of the streptococcal M protein. It is encoded by a gene- tically  
 CC engineered gene introduced into the genome of a vaccinia or fowlpox  
 CC virus. The resultant DNA complex is useful as a vaccine for  
 CC immunoprotection against streptococcal infections. The M'6 polypeptide is  
 CC the part of protein M responsible for virulence. (Updated on 25-MAR-2003  
 CC to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 234 AA;

Query Match 100.0%; Score 64; DB 2; Length 234;  
 Best Local Similarity 100.0%; Pred. No. 0.026;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKALE 14  
 |||||  
 Db 85 ASREAKQVEKALE 98

RESULT 10  
 AAB03118  
 ID AAB03118 standard; protein; 236 AA.

XX  
 AC AAB03118;

DT 10-OCT-2000 (first entry)

DE C-terminus of Streptococcus pyogenes M protein M5.

XX Multivalent hybrid M protein; M5; group A streptococcus; serotype;

XX immunogenic; sero-specific antibody; streptococcal infection;

XX cross reactivity; vaccine; acute rheumatic fever; ARF;

XX rheumatic heart disease; streptococcal pharyngitis; strep throat;

XX pneumonia.

OS Streptococcus pyogenes.

PN US6063386-A.

XX 16-MAY-2000.

PF 15-SEP-1997; 97US-00937271.

PR 16-SEP-1992; 92US-00945954.

XX (UYTE-) UNIV TENNESSEE RES CORP.

XX Lederer JW, Dale JB;

XX WPI; 2000-364475/31.

PT New immunogenic recombinant hybrid M protein comprising amino-terminal  
 PT peptide fragments of Streptococcal M protein useful as vaccine against  
 PT rheumatic fever and infections leading to rheumatic fever.

XX Disclosure; Col 45-46; 62pp; English.

CC The invention relates to multivalent immunogenic hybrid group A  
 CC streptococcal M proteins comprising N-terminal peptide fragments of M  
 CC proteins that elicit opsonic antibodies against multiple serotypes of  
 CC group A streptococci (e.g., Streptococcus pyogenes). The antibodies  
 CC generated using the hybrid proteins are against one or more M protein  
 CC serotypes, where at least one of the serotypes is M1, M2, M18 or M19. The  
 CC invention also encompasses a recombinant DNA molecule comprising a  
 CC nucleotide sequence that encodes a multivalent hybrid M protein; and a  
 CC method for immunizing a mammal against streptococcal infections,  
 CC comprising administering an immunogenic multivalent hybrid M protein to  
 CC the mammal. The multivalent hybrid M proteins are useful for eliciting  
 CC opsonic or protective antibodies to the M proteins of different serotypes  
 CC of group A streptococci, and may therefore be used as vaccines to protect  
 CC against and control infection by type A streptococci. Type A streptococci  
 CC are not only responsible for streptococcal pharyngitis (strep throat),  
 CC forms of pneumonia and a condition resembling toxic shock, but are also  
 CC involved in the development of acute rheumatic fever (ARF) and rheumatic  
 CC heart disease. In a patient with ARF, antibodies formed during a group A  
 CC streptococcal infection are also cross-reactive with heart tissue, which  
 CC indicates that the streptococci and host tissue contain similar antigenic  
 CC motifs. The new multivalent vaccines are capable of raising sero-specific  
 CC antibodies against various serotypes of group A streptococci which are  
 CC not cross-reactive with human heart tissue. The present sequence  
 CC represents the C-terminal half of the Streptococcus pyogenes M5 protein,  
 CC which is a component of the tetravalent M24-M5-M6-M19 hybrid M protein,  
 CC AAB03117

SQ Sequence 236 AA;

Query Match 100.0%; Score 64; DB 3; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 0.026;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERALE 14  
 |||||  
 DB 87 ASREAKKOVERALE 100

## RESULT 11

AAR50228 standard; protein; 254 AA.

AAAR50228;

AC 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 09-OCT-1994 (first entry)

DE Sequence of a fragment of Group A streptococcal surface protein M5 and a carrier of the COOH-terminal portion of M5.

KM B subunit; labile toxin; M protein; fusion protein; antigen;  
 KM Group A streptococci; rheumatic fever; pharyngitis.

OS Streptococcus sp.

PN W09406465-A1.

PD 31-MAR-1994.

PF 15-SEP-1993; 93WO-US008704.

PR 16-SEP-1992; 92US-00945860.

PA (UYTE-) UNIV TENNESSEE RES CORP.

PI Dale JB;

DR WPI; 1994-118162/14.

DR N-PSDB; AAQ45160.

PT New recombinant hybrid streptococcal M protein antigen(s) - which elicit opsonic antibodies without eliciting cross-reactive antibodies to mammalian heart tissue.

PS Disclosure; Fig 4; 45PP; English.

CC The surface M protein of Group A streptococci is the major virulence factor and protective antigen of these organisms. However, there are a tremendous number of M protein serotypes. The invention provides a recombinant M protein antigens comprising a gene encoding a carrier

CC protein and an NH2 or COOH terminal M protein fragment carrying one or more epitopes. The carrier may be the B subunit of E.coli labile toxin

CC (LT-B) or the C-repeat portion of a streptococcal M protein. For example, CC AAQ45160/R50228 comprises an antigen which is an M5 hapten fragment of 16

CC AA joined by a BamHI restriction site to a carrier which is the COOH-terminal half of M5. The carrier includes 2.5 C-repeats, which each commence with the tetrapeptide NKIS. A linker could be inserted at the BamHI site. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)

SO Sequence 254 AA;

QY Query Match 100.0%; Score 64; DB 2; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 0.028;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERALE 14  
 |||||

DB 105 ASREAKKOVERALE 118

## RESULT 12

AAR20128 standard; protein; 281 AA.

ID AAR20128

XX

AC AAR20128;

DT 27-AUG-2003 (revised)  
 DT 15-APR-1992 (first entry)

DE Sequence encoded by truncated M1 gene.

KM Protein H; immunoglobulin G; IgG; antibody; autoimmune disease.

OS Streptococcus sp.

FT Key Location/Qualifiers

FT Region 1..28 /label= p16M1

FT Region 29..70 /label= C1

FT Region 71..112 /label= C2

FT Region 113..155 /label= C3

FT Region 156..176 /label= C4

FT Region 177..281 /label= D

PN W09119740-A.

PD 26-DEC-1991.

PF 21-JUN-1990; 90SE-00002212.

PR 21-JUN-1990; 90SE-00002212.

PA (HIGH-) HIGHTECH RECEPTOR A.

PI Schmidt KH, Akesson P, Cooney J, Bjorck L;

DR WPI; 1992-024366/03.

DR N-PSDB; AAQ20292.

PT New IgG binding proteins H' lacking an albumin binding sequence - useful in purification of excess IgG from blood and to diagnose autoimmune diseases.

PS Disclosure; Fig 8; 37PP; English.

CC The inventors claim a protein prod. by a strain of Gp.A Streptococci. The protein has the AA sequence of protein H but lacks at least some part of the C and D regions (responsible for binding albumin), esp. it lacks the

CC whole of these regions and extends for AA1 to AA158. Compared with CC natural protein H, it is more specific and may be used as part of a kit for the binding, separation and identification of human IgG. The same

CC sequences appear in W09119741. (Updated on 27-AUG-2003 to correct OS field.)

SO Sequence 281 AA;

QY Query Match 100.0%; Score 64; DB 2; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERALE 14  
 |||||

DB 132 ASREAKKOVERALE 145

## RESULT 13

AAR50229 standard; protein; 284 AA.

ID AAR50229

AC AAR50229;

DT 27-AUG-2003 (revised)



DT 10-OCT-2000 (first entry)  
 XX S. pyogenes hybrid M protein (M24-M5-M6-M19), SEQ ID NO:10.  
 XX  
 DE Multivalent hybrid M protein; group A streptococcus; serotype;  
 KM immunogenic; sero-specific antibody; streptococcal infection;  
 KM cross reactivity; vaccine; acute rheumatic fever; ARF;  
 KM rheumatic heart disease; streptococcal pharyngitis; strep throat;  
 KM pneumonia.  
 XX  
 OS Streptococcus pyogenes.  
 OS Synthetic.  
 OS  
 PN US6063386-A.  
 XX  
 PD 16-MAY-2000.  
 XX  
 PF 15-SEP-1997; 97US-00937271.  
 XX  
 PR 16-SEP-1992; 92US-00945954.  
 XX  
 PA (UYTE-) UNIV TENNESSEE RES CORP.  
 XX  
 PI Lederer JW, Dale JB;  
 XX  
 DR WPI; 2000-364475/31.  
 XX  
 DR N-PSDB; AAA57897.  
 XX  
 PT New immunogenic recombinant hybrid M protein comprising amino-terminal  
 PT peptide fragments of streptococcal M protein useful as vaccine against  
 PT rheumatic fever and infections leading to rheumatic fever.  
 XX  
 PS Disclosure; Fig 8A-B; 62pp; English.  
 XX  
 CC The invention relates to multivalent immunogenic hybrid group A  
 CC streptococcal M proteins comprising N-terminal peptide fragments of M  
 CC proteins that elicit opsonic antibodies against multiple serotypes of  
 CC group A streptococci (e.g., Streptococcus pyogenes). The antibodies  
 CC generated using the hybrid proteins are against one or more M protein  
 CC serotypes, where at least one of the serotypes is M1, M3, M18 or M19. The  
 CC invention also encompasses a recombinant DNA molecule comprising a  
 CC nucleotide sequence that encodes a multivalent hybrid M protein; and a  
 CC method for immunizing a mammal against streptococcal infections,  
 CC comprising administering an immunogenic multivalent hybrid M protein to  
 CC the mammal. The multivalent hybrid M proteins are useful for eliciting  
 CC opsonic or protective antibodies to the M proteins of different serotypes  
 CC of group A streptococci, and may therefore be used as vaccines to protect  
 CC against and control infection by type A streptococci. Type A streptococci  
 CC are not only responsible for streptococcal pharyngitis (strep throat),  
 CC forms of pneumonia and a condition resembling toxic shock, but are also  
 CC involved in the development of acute rheumatic fever (ARF) and rheumatic  
 CC heart disease. In a patient with ARF, antibodies formed during a group A  
 CC streptococcal infection are also cross-reactive with heart tissue, which  
 CC indicates that the streptococci and host tissue contain similar antigenic  
 CC motifs. The new multivalent vaccines are capable of raising sero-specific  
 CC antibodies against various serotypes of group A streptococci which are  
 CC not cross-reactive with human heart tissue. Sequences AAB03113-B03117,  
 CC AAB03119-B03121 and AAB03123-B03124 represent multivalent hybrid  
 CC Streptococcus pyogenes M proteins generated in the disclosure of the  
 CC invention.  
 CC  
 XX  
 SQ Sequence 305 AA;  
 Query Match 100.0%; Score 64; DB 3; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 156 ASREAKKQVEKALE 169  
 RESULT 16

AAP90955  
 ID AAP90955 standard; protein; 441 AA.  
 XX  
 AC AAP90955;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 23-FEB-1990 (first entry)  
 XX  
 DE M6 streptococcal protein.  
 XX  
 KM Immunoglobulin.  
 KM  
 KW Streptococcus sp; 'group A'.  
 OS  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Region 216..235  
 FT Region 248..269  
 FT Region 275..284  
 XX  
 XX WO8909064-A.  
 XX  
 XX  
 PD 05-OCT-1989.  
 XX  
 PF 13-MAR-1989; 89NO-US001026.  
 XX  
 PR 25-MAR-1988; 88US-00173380.  
 PR 27-FEB-1989; 89US-00315588.  
 XX  
 PA (UYRQ ) UNIV ROCKEFELLER.  
 XX  
 PI Fischetti VA;  
 XX  
 DR WPI; 1989-309382/42.  
 XX  
 PT Polypeptide(s) from streptococcal M protein - used to prepare vaccines  
 PT for providing protection against streptococcal infection.  
 XX  
 PS Disclosure; Fig 1; 22pp; English.  
 XX  
 CC The regions (pref. conjugated to a natural carrier, eg cholera toxin)  
 CC above can elicit an secretory Ig response in a mammal. They are used in  
 CC vaccines against streptococcal infection, and give protection to  
 CC different serotypes. (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to  
 CC correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 441 AA;  
 Query Match 100.0%; Score 64; DB 1; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 0.051;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 292 ASREAKKQVEKALE 305  
 RESULT 17  
 AAR41780  
 ID AAR41780 standard; protein; 441 AA.  
 XX  
 AC AAR41780;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 25-MAR-1994 (first entry)  
 XX  
 DE Streptococcus pyogenes M6 protein.  
 XX  
 KM Group A Streptococci; coiled coil surface antigen; anchor region;  
 KM gram positive bacterium; M protein; fusion protein; immunogen.  
 XX  
 OS Streptococcus pyogenes.

```

XX Key Location/Qualifiers
FH Region 1..362
FT /label= extracellular region
FT /note= "coiled structure"
FT Region 298..441
FT /note= "cell-associated region"
FT Region 298..370
FT /label= carbohydrate_segment
FT Region 371..416
FT /label= peptidoglycan_segment
FT Region 407..441
FT /label= anchor
FT /note= "preferred anchor region for use in novel fusion
FT proteins; includes a 3 amino acid spacer between the
FT consensus anchor motif and the hydrophobic segment"
FT Region 407..412
FT /note= "consensus LPSTGB motif common to anchor regions
FT of gram-positive bacteria"
FT Region 417..435
FT /label= hydrophobic_segment
FT Region 436..441
FT /label= tail_segment
FT /note= "highly charged"
FT
FT
XX MO9318163-A2.
XX
XX 16-SEP-1993.
XX
XX 12-MAR-1993; 93WO-US002355.
XX
XX 13-MAR-1992; 92US-00851082.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Fischetti VA, Pozzi G, Schneewind O,
XX WPI; 1993-303477/38.
XX
XX
XX Gene encoding hybrid surface protein of gram positive bacteria - useful
XX for preparing vaccine compen. for protecting animals from bacterial
XX infection.
XX
XX Disclosure and Claim 14; Fig 3; 85pp; English.
XX
XX Hybrid surface proteins are claimed in which an active polypeptide (e.g.
XX a surface antigen from a mammalian tumour cell, sperm or an allergen,
XX bacterium, virus, parasite or fungus) is fused to an anchor region from a
XX surface antigen normally expressed on the cell surface of gram positive
XX bacteria. The anchor segment is pref. derived from a streptococcal M
XX protein. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 441 AA;
SQ
Query Match 100.0%; Score 64; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASREAKKQVEKALE 14
Db 292 ASREAKKQVEKALE 305

```

## RESULT 18

```

ADG62862 ID ADG62862 standard; protein; 441 AA.
XX
XX ADG62862;
XX
XX 11-MAR-2004 (first entry)
XX
XX Streptococcus sp. M6 protein.
XX

```

```

KW Antigen conjugate; immune response; infection; C repeat region;
KW cholera toxin B; CTB.
XX
XX Streptococcus sp.
XX
XX US6602507-B1.
XX
XX 05-AUG-2003.
XX
XX 06-JAN-1995; 95US-00369295.
XX
XX 25-MAR-1988; 88US-00173380.
XX
XX 27-FEB-1989; 89US-00315588.
XX
XX 19-JUN-1990; 90US-00540101.
XX
XX 03-MAR-1992; 92US-00845865.
XX
XX 28-MAY-1993; 93US-00068598.
XX
XX 22-FEB-1994; 94US-00200913.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX
XX Fischetti VA;
XX
XX WPI; 2003-895270/82.
XX
XX
XX New antigen conjugates comprising polypeptides from the conserved exposed
XX region of group A streptococcal M protein, useful for preparing vaccines
XX for controlling streptococcal infection a mammal.
XX
XX
XX Disclosure; Fig 1; 12pp; English.
XX
XX
XX The invention relates an antigen conjugate comprising a linkable carrier
XX covalently bound to a polypeptide consisting of five or more amino acid
XX residues from the conserved exposed region of group A streptococcal M
XX protein. The antigen conjugate elicits a protective immune response to
XX streptococcal infection in a mammal when administered mucosally. The
XX antigen conjugates and polypeptides are useful for preparing vaccines for
XX controlling streptococcal infection or protecting against streptococcal
XX infection in a mammal. The polypeptides were all within the C repeat
XX region of the M protein of type 6 Streptococci (strain 471). The
XX polypeptides were purified by high performance liquid chromatography.
XX Highly purified cholera toxin B (CTB) was derivatised at primary amino
XX groups with the heterobifunctional cross-linking agent N'-succinimidyl 3-
XX (2-pyridyldithio) propionate (SPDP). The synthetic peptides were mixed
XX with SPDP-derivatised CTB at a 1.5:1 ratio. An average of 1.06-1.27
XX peptide molecules were covalently linked per CTB monomer. The present
XX sequence is M6 protein from strain D471 of group A Streptococcus sp.
XX
XX Sequence 441 AA;
SQ
Query Match 100.0%; Score 64; DB 7; Length 441;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASREAKKQVEKALE 14
Db 292 ASREAKKQVEKALE 305

```

## RESULT 19

```

AAW08927 ID AAW08927 standard; protein; 483 AA.
XX
XX AAW08927;
XX
XX 31-MAR-1997 (first entry)
XX
XX
XX Type-6 M-protein.
XX
XX Type-6 M-protein; emm6 gene; Streptococcus pyogenes; signal peptide;
XX anchor peptide; C-terminal sorting signal; surface display;
XX fusion protein; protease-deficient; Streptococcus gordonii; SPFX vector;
XX fusion protein cleavage; spacer; protein secretion; TEV-Nia protease;
XX diagnostic; recombinant vaccine; therapy.
XX

```

```

XX OS Streptococcus pyogenes.
XX Key Location/Qualifiers
XX Key 1..58
XX Peptide /note= "Preferred N-terminal secretion sequence (claim
XX 23)"
XX Peptide 1..42
XX Peptide /note= "Signal peptide (claim 22)"
XX Peptide 263..342
XX Peptide /note= "Preferred spacer peptide (claim 32)"
XX Peptide 344..483
XX Peptide /note= "Preferred C-terminal anchoring sequence (claim
XX 25)"
XX Domain 405..458
XX Domain /note= "Proline/glycine-rich cell wall domain"
XX Peptide 449..454
XX Peptide /note= "Conserved C-terminal motif (AAW08924, claim 9)"
XX Domain 459..477
XX Domain /note= "Hydrophobic membrane-spanning domain"
XX Domain 478..483
XX Domain /note= "C-terminal charged tail"
XX PN W09640943-A1.
XX PD 19-DEC-1996.
XX PP 06-JUN-1996; 96WO-US009965.
XX PR 07-JUN-1995; 95US-00472244.
XX PA (UYRQ ) UNIV ROCKEFELLER.
XX PI Darzins A, Whitehead S, Hraby D;
XX DR WPI: 1997-052336/05.
XX DR N-PSDB; AAT49317.
XX FT Expressing heterologous proteins in gram-positive bacteria - produces
XX PT proteins anchored to host's cell surface, or secreted from cell.
XX PS Disclosure; Fig 3; 41pp; English.
XX CC This sequence represents Streptococcus pyogenes type-6 M-protein, which
XX CC is a cell wall surface protein. The N-terminal signal peptide, part of
XX CC the mature N-terminal sequence, the C-terminal sorting signal (anchor)
XX CC sequence and optionally a spacer region from the protein may be expressed
XX CC as a fusion with a target protein in a protease-deficient Gram-positive
XX CC host, e.g. the human oral commensal Streptococcus gordonii, in a new
XX CC cloning method involving the SPEX vector series (e.g. plasmid pSPEX1A).
XX CC The C-terminal sorting signal contains a proline/glycine-rich region
XX CC (with a conserved cell wall anchoring motif, AAW08924), a hydrophobic
XX CC membrane-spanning region and a charged tail. The resulting fusion protein
XX CC is expressed on the cell surface, and may be cleaved with e.g. tobacco-
XX CC etch virus TEV-N1a protease (at a site adjacent to the anchor, e.g.
XX CC AAW08925-26) for sample purification. A spacer region from M-protein may
XX CC be included to reduce steric hindrance during fusion protein cleavage. A
XX CC vector without the anchor sequence may be used for protein secretion. The
XX CC method may be used in production of recombinant diagnostic, vaccine and
XX CC therapeutic proteins
XX SQ Sequence 483 AA;
XX
XX Query Match 100.0%; Score 64; DB 2; Length 483;
XX Best Local Similarity 100.0%; Pred. No. 0.056;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ASREAKQVEKALE 14
XX |||||||||||
XX Db 334 ASREAKQVEKALE 347
XX
XX RESULT 20

```

```

ABP30015
XX ID ABP30015 standard; protein; 484 AA.
XX AC ABP30015;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 9206.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus pyogenes.
XX PN W0200234771-A2.
XX PD 02-MAY-2002.
XX PP 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
XX PI Tetelini H;
XX DR WPI; 2002-352536/38.
XX DR N-PSDB; ABN70646.
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 1; Page 4044; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (I), nucleic acids encoding (II), ABN6044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (II) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassays, and distinguishing/identifying
XX CC Streptococcus proteins
XX SQ Sequence 484 AA;
XX
XX Query Match 100.0%; Score 64; DB 5; Length 484;
XX Best Local Similarity 100.0%; Pred. No. 0.056;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ASREAKQVEKALE 14
XX |||||||||||
XX Db 335 ASREAKQVEKALE 348
XX
XX RESULT 21
XX ADR83969
XX ID ADR83969 standard; protein; 484 AA.
XX
XX

```

[illegible]

PN	WO2004046380-A1.
PD	
XX	03-JUN-2004.
XX	
PF	19-NOV-2003; 2003WO-GB005011.
PR	20-NOV-2002; 2002GB-00027026.
XX	
PA	(MOLE-) MOLECULAR SKINCARE LTD.
XX	
P1	Cork MJ, Ward SJ, Tazi-Ahmini R;
XX	
DR	WPI; 2004-449755/42.
XX	
PT	Diagnosing or detecting polymorphism in alpha-helix coiled-coil rod
PT	homolog (HCR) gene in human, by determining sequence of HCR gene in
PT	polymorphic position and determining status of human by reference to
PT	polymorphism in HCR.
XX	
PS	Claim 14; Page 48-50; 55pp; English.
XX	
CC	The present invention relates to a method of diagnosing or detecting a
CC	polymorphism in the alpha-helix coiled-coil rod homologue (HCR) gene in a
CC	human, which involves determining the sequence of the human HCR gene in
CC	at least one polymorphic position of HCR and determining the status of
CC	the human by reference to the polymorphism in HCR. The method is useful
CC	for diagnosing or detecting a polymorphism in HCR gene in a human, which
CC	can in turn be used for developing a diagnostic assay for susceptibility
CC	and/or predisposition to guttate psoriasis or chronic plaque psoriasis.
CC	Polymorphisms diagnosed or detected using the method are useful for
CC	developing a medicament for treating guttate psoriasis or psoriasis
CC	vulgaris, where the medicament interferes with the antigenic properties
CC	of the polymorphic protein. The present sequence is the Streptococcus
CC	pyogenes M21 protein, which can be used as a target for the medicament
CC	described in the invention.
XX	
SQ	Sequence 539 AA;
	Query Match 100.0%; Score 64; DB 8; Length 539;
	Best Local Similarity 100.0%; Pred. No. 0.063;
	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ASRAKQVEKALE 14
DB	390 ASRAKQVEKALE 403
RESULT 23	
AAW04353	
ID	AAW04353 standard; protein; 28 AA.
XX	
AC	AAW04353;
XX	
DT	02-DEC-1996 (first entry)
XX	
DE	Chimaeric peptide (J7) contg. Streptococcal M protein peptide p145.
XX	
KM	Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;
KM	conformational epitope; alpha-helix; GCN4; leucine zipper; detection;
KM	mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;
XX	diagnosis.
XX	
OS	Synthetic.
XX	
XX	
FH	Key Location/Qualifiers
FT	Peptide 15..26
FT	/note="p145 conformational B-cell epitope"
XX	
XX	WO9611944-A1.
PN	
PD	25-APR-1996.
XX	
PD	16-OCT-1995; 95WO-AU000681.

```

XX 14-OCT-1994; 94AU-00008851.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA (UYME ) UNIV MELBOURNE.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.
PA (CSLC-) CSL LTD.
XX
PI Cooper JA, Relf WA, Good MF, Saul AJ;
XX WPI; 1996-221939/22.
DR
XX
XX New chimeric peptide(s) including a conformational epitope - inserted
PT into a peptide having similar native conformation, useful in vaccines and
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.
XX
XX Example 12; Fig 1C; 99pp; English.
PS
XX
XX The present peptide is a chimeric peptide (CP) contg. the Streptococcal
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735
CC (1992)). The CP comprises a B-cell conformational epitope from within
CC p145 inserted into a 2nd peptide, pref. an alpha-helical coil based on
CC the GCN4 leucine zipper peptide (AA97395). The 2nd peptide has a similar
CC conformation, enabling the epitope to be presented in an immunologically
CC active conformation. The CP can be used in a novel detection/mapping
CC process, e.g. to determine the min. epitope required to induce opsonic
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised
CC against the CP can be used for immunotherapy and diagnosis, while the CP
CC can be used diagnostically to detect Ab. The reactivity of the present
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145
XX
SQ Sequence 28 AA;
Query Match 89.1%; Score 57; DB 2; Length 28;
Best Local Similarity 85.7%; Pred. No. 0.038;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASREAKKQVEKALE 14
Db 9 ASREAKKQVEKAVK 22

```

```

PA (UYME ) UNIV MELBOURNE.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.
PA (CSLC-) CSL LTD.
XX
PI Cooper JA, Relf WA, Good MF, Saul AJ;
XX WPI; 1996-221939/22.
DR
XX
XX New chimeric peptide(s) including a conformational epitope - inserted
PT into a peptide having similar native conformation, useful in vaccines and
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.
XX
XX Example 13; Page 37; 99pp; English.
PS
XX
XX The present peptide is a fragment of the Streptococcal M protein peptide
CC p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735 (1992)), used in the
CC construction of a novel chimeric peptide (CP). The CP comprises a B-cell
CC conformational epitope from within p145, inserted into a 2nd peptide,
CC pref. an alpha-helical coil based on the GCN4 leucine zipper peptide
CC (AA97395). The 2nd peptide has a similar conformation, enabling the
CC epitope to be presented in an immunologically active conformation. The CP
CC can be used in a novel detection/mapping process, e.g. to determine the
CC min. epitope required to induce opsonic antibodies (Ab), and in vaccines
CC against gp. A Streptococci. Ab raised against the CP can be used for
CC immunotherapy and diagnosis, while the CP can be used diagnostically to
CC detect Ab. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 12 AA;
Query Match 87.5%; Score 56; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 REAKKQVEKALE 14
Db 1 REAKKQVEKALE 12

```

```

RESULT 24
AA97415
ID AA97415 standard; peptide; 12 AA.
XX
AC AA97415;
XX
DT 27-AUG-2003 (revised)
DT 02-DEC-1996 (first entry)
XX
DE Streptococcal M protein peptide, p145, fragment U(1)9.
XX
KM Streptococcal; M protein; peptide; p145; chimeric; chimeric; B-cell;
KM conformational epitope; alpha-helix; GCN4; leucine zipper; detection;
KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;
KM diagnosis.
XX
OS Streptococcus sp.
XX
PN MO9611944-A1.
XX
PD 25-APR-1996.
XX
PF 16-OCT-1995; 95WO-AU000681.
XX
PR 14-OCT-1994; 94AU-00008851.
XX
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.

```

```

RESULT 25
AAW04354
ID AAW04354 standard; protein; 28 AA.
XX
AC AAW04354;
XX
DT 02-DEC-1996 (first entry)
XX
DE Chimeric peptide (U8) contg. Streptococcal M protein peptide p145.
XX
KM Streptococcal; M protein; peptide; p145; chimeric; chimeric; B-cell;
KM conformational epitope; alpha-helix; GCN4; leucine zipper; detection;
KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;
KM diagnosis.
XX
OS Synthetic.
XX
FT Key Location/Qualifiers
FT Peptide 16..27
FT /note= "p145 conformational B-cell epitope"
XX
PN MO9611944-A1.
XX
PD 25-APR-1996.
XX
PF 16-OCT-1995; 95WO-AU000681.
XX
PR 14-OCT-1994; 94AU-00008851.
XX
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA (UYME ) UNIV MELBOURNE.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.

```



PA (CSLC-) CSL LTD.  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX WPI; 1996-221939/22.  
DR  
XX  
XX  
PT New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 12; Fig 1C; 99pp; English.  
XX  
XX The present peptide is a chimaeric peptide (CP) contg. the Streptococcal  
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
CC (1992)). The CP comprises a B-cell conformational epitope from within  
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against sp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
CC  
XX  
SQ Sequence 28 AA;  
Query Match 87.5%; Score 56; DB 2; Length 28;  
Best Local Similarity 92.3%; Pred. No. 0.054; Mismatches 0; Indels 0; Gaps 0;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SREAKQVEKALE 14  
Db 9 SREAKQVEKALK 21  
RESULT 26  
AAW04355  
ID AAW04355 standard; protein; 28 AA.  
XX  
AC AAW04355;  
XX  
DT 02-DEC-1996 (first entry)  
XX  
DE Chimaeric peptide (J9) contg. Streptococcal M protein peptide p145.  
XX  
XX Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
XX diagnosis.  
XX  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Peptide 17.28  
FT /note="p145 conformational B-cell epitope"  
XX  
XX W09611944-A1.  
PN  
XX  
XX 25-APR-1996.  
PD  
XX 16-OCT-1995; 95WO-AU000681.  
PF  
XX 14-OCT-1994; 94AU-00008851.  
PR  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX

PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX WPI; 1996-221939/22.  
DR  
XX  
XX  
PT New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 12; Fig 1C; 99pp; English.  
XX  
XX The present peptide is a chimaeric peptide (CP) contg. the Streptococcal  
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
CC (1992)). The CP comprises a B-cell conformational epitope from within  
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against sp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
CC  
XX  
SQ Sequence 28 AA;  
Query Match 87.5%; Score 56; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.054; Mismatches 0; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 REAKQVEKALE 14  
Db 9 REAKQVEKALK 20  
RESULT 27  
AAR97413  
ID AAR97413 standard; peptide; 12 AA.  
XX  
AC AAR97413;  
XX  
DT 27-AUG-2003 (revised)  
DT 02-DEC-1996 (first entry)  
XX  
DE Streptococcal M protein peptide, p145, fragment J(1)7.  
XX  
XX Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
XX diagnosis.  
XX  
OS Streptococcus sp.  
XX  
XX W09611944-A1.  
PN  
XX  
XX 25-APR-1996.  
PD  
XX 16-OCT-1995; 95WO-AU000681.  
PF  
XX 14-OCT-1994; 94AU-00008851.  
PR  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
XX Cooper JA, Relf WA, Good MF, Saul AJ;  
XX WPI; 1996-221939/22.  
DR  
XX New chimeric peptide(s) including a conformational epitope - inserted  
PT

PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determin. of minimal epitope(s) or for mapping amphipathic helices.

XX Example 13; Page 37; 99pp; English.

CC The present peptide is a fragment of the Streptococcal M protein peptide  
 CC p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735 (1992)), used in the  
 CC construction of a novel chimeric peptide (CP). The CP comprises a B-cell  
 CC conformational epitope from within p145, inserted into a 2nd peptide,  
 CC pref. an alpha-helical coil based on the GCN4 leucine zipper peptide  
 CC (AAR97395). The 2nd peptide has a similar conformation, enabling the  
 CC epitope to be presented in an immunologically active conformation. The CP  
 CC can be used in a novel detection/mapping process, e.g. to determine the  
 CC min. epitope required to induce opsonic antibodies (Ab), and in vaccines  
 CC against gp. A Streptococci. Ab raised against the CP can be used for  
 CC immunotherapy and diagnosis, while the CP can be used diagnostically to  
 CC detect Ab. (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 12 AA;

Query Match 85.9%; Score 55; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASREAKQVEKA 12  
 |||||  
 Db 1 ASREAKQVEKA 12

RESULT 28

AAR97414 ID AAR97414 standard; peptide; 12 AA.

XX AAR97414;

DT 27-AUG-2003 (revised)  
 DT 02-DEC-1996 (first entry)

XX Streptococcal M protein peptide, p145, fragment J(1)8.

KM Streptococcal; M protein; peptide: p145; chimeric; chimeric; B-cell;  
 KM conformational epitope: alpha-helix; GCN4; leucine zipper; detection;  
 KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
 KM diagnosis.

XX Streptococcus sp.

PN W09611944-A1.

XX 25-APR-1996.

PF 16-OCT-1995; 95WO-AU000681.

PR 14-OCT-1994; 94AU-00008851.

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.

PA (UYME ) UNIV MELBOURNE.

PA (HALT-) HALT. INST MEDICAL RES WALTER & ELIZA.

PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.

PI (GSLC-) CSL LTD.

PI Cooper JA, Relf WA, Good MF, Saul AJ;

DR WPI; 1996-221939/22.

XX New chimeric peptide(s) including a conformational epitope - inserted  
 PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
 XX Example 13; Page 37; 99pp; English.  
 XX The present peptide is a fragment of the Streptococcal M protein peptide

CC p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735 (1992)), used in the  
 CC construction of a novel chimeric peptide (CP). The CP comprises a B-cell  
 CC conformational epitope from within p145, inserted into a 2nd peptide,  
 CC pref. an alpha-helical coil based on the GCN4 leucine zipper peptide  
 CC (AAR97395). The 2nd peptide has a similar conformation, enabling the  
 CC epitope to be presented in an immunologically active conformation. The CP  
 CC can be used in a novel detection/mapping process, e.g. to determine the  
 CC min. epitope required to induce opsonic antibodies (Ab), and in vaccines  
 CC against gp. A Streptococci. Ab raised against the CP can be used for  
 CC immunotherapy and diagnosis, while the CP can be used diagnostically to  
 CC detect Ab. (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 12 AA;

Query Match 85.9%; Score 55; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SREAKQVEKAL 13  
 |||||  
 Db 1 SREAKQVEKAL 12

RESULT 29

ADG62865 ID ADG62865 standard; peptide; 22 AA.

XX ADG62865;

DT 11-MAR-2004 (first entry)

XX Streptococcus sp. M6 protein peptide epitope (248-269).

KM Antigen conjugate; immune response; infection; C repeat region;  
 KM cholera toxin B; CTB.

XX Streptococcus sp.

PN US6602507-B1.

PD 05-AUG-2003.

PF 06-JAN-1995; 95US-00369295.

PR 25-MAR-1988; 88US-00173380.

PR 27-FEB-1989; 89US-00315588.

PR 19-JUN-1990; 90US-00540101.

PR 03-MAR-1992; 92US-00845865.

PR 28-MAY-1993; 93US-00068598.

PR 22-FEB-1994; 94US-00200913.

PA (UYRQ ) UNIV ROCKEFELLER.

PI Fischetti VA;

DR WPI; 2003-895270/82.

XX New antigen conjugates comprising polypeptides from the conserved exposed  
 PT region of group A streptococcal M protein, useful for preparing vaccines  
 PT for controlling streptococcal infection a mammal.

XX Claim 5; Col 13; 12pp; English.

XX The invention relates an antigen conjugate comprising a linkable carrier  
 CC covalently bound to a polypeptide consisting of five or more amino acid  
 CC residues from the conserved exposed region of group A streptococcal M  
 CC protein. The antigen conjugate elicits a protective immune response to  
 CC streptococcal infection in a mammal when administered mucosally. The  
 CC antigen conjugates and polypeptides are useful for preparing vaccines for  
 CC controlling streptococcal infection or protecting against streptococcal  
 CC infection in a mammal. The polypeptides were all within the C repeat  
 CC region of the M protein of type 6 Streptococci (strain 471). The  
 CC polypeptides were purified by high performance liquid chromatography.

CC Highly purified cholera toxin B (CTB) was derivatised at primary amino  
 CC groups with the heterobifunctional cross-linking agent N'-succinimidyl 3-  
 CC (2-pyridyldithio) propionate (SPDP). The synthetic peptides were mixed  
 CC with SPDP-derivatised CTB at a 1.5:1 ratio. An average of 1.06-1.27  
 CC peptide molecules were covalently linked per CTB monomer. The present  
 CC sequence is M6 protein peptide epitope from strain D471 of group A  
 CC Streptococcus sp.

SQ Sequence 22 AA;

Query Match 82.8%; Score 53; DB 7; Length 22;  
 Best Local Similarity 92.3%; Pred. No. 0.13;  
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKAL 13  
 |||||  
 Db 3 ASREAKKQVEKDL 15

RESULT 30  
 ADG62863  
 ID ADG62863 standard; peptide; 65 AA.

AC ADG62863;

DT 11-MAR-2004 (first entry)

DE Streptococcus sp. M6 protein C repeat region peptide.

KM Antigen conjugate; immune response; infection; C repeat region;

KW cholera toxin B; CTB.

OS Streptococcus sp.

PH Key Location/Qualifiers

FT Region 1..23

FT /note= "C1 block"

FT Region 24..42

FT /note= "Spacer"

FT Region 43..65

FT /note= "C2 block"

PN US6602507-B1.

PD 05-AUG-2003.

PF 06-JAN-1995; 95US-00369295.

PR 25-MAR-1988; 88US-00173380.

PR 27-FEB-1989; 89US-00315588.

PR 19-JUN-1990; 90US-00540101.

PR 03-MAR-1992; 92US-00845865.

PR 28-MAY-1993; 93US-00068598.

PR 22-FEB-1994; 94US-00200913.

PA (UYRO ) UNIV ROCKEFELLER.

PI Fleichetti VA;

DR WPI; 2003-895270/82.

PT New antigen conjugates comprising polypeptides from the conserved exposed  
 PT region of group A streptococcal M protein, useful for preparing vaccines  
 PT for controlling streptococcal infection a mammal.

PS Disclosure; Fig 3, 12pp; English.

XX The invention relates an antigen conjugate comprising a linkable carrier  
 CC covalently bound to a polypeptide consisting of five or more amino acid  
 CC residues from the conserved exposed region of group A streptococcal M  
 CC protein. The antigen conjugate elicits a protective immune response to  
 CC streptococcal infection in a mammal when administered mucosally. The  
 CC antigen conjugates and polypeptides are useful for preparing vaccines for

CC controlling streptococcal infection or protecting against streptococcal  
 CC infection in a mammal. The polypeptides were all within the C repeat  
 CC region of the M protein of type 6 Streptococci (strain 471). The  
 CC polypeptides were purified by high performance liquid chromatography.

CC Highly purified cholera toxin B (CTB) was derivatised at primary amino  
 CC groups with the heterobifunctional cross-linking agent N'-succinimidyl 3-  
 CC (2-pyridyldithio) propionate (SPDP). The synthetic peptides were mixed  
 CC with SPDP-derivatised CTB at a 1.5:1 ratio. An average of 1.06-1.27  
 CC peptide molecules were covalently linked per CTB monomer. The present  
 CC sequence is M6 protein C repeat region from strain D471 of group A  
 CC Streptococcus sp.

SQ Sequence 65 AA;

Query Match 82.8%; Score 53; DB 7; Length 65;  
 Best Local Similarity 92.3%; Pred. No. 0.4;  
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKAL 13  
 |||||  
 Db 16 ASREAKKQVEKDL 28

RESULT 31  
 AAW04352

ID AAW04352 standard; protein; 28 AA.

AC AAW04352;

DT 02-DEC-1996 (first entry)

DE Chimaeric peptide (v6) contg. Streptococcal M protein peptide p145.

DX

KM Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;

KW conformational epitope; alpha-helix; GCM4; leucine zipper; detection;

KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;

KW diagnosis.

OS Synthetic.

PH Key Location/Qualifiers

FT Peptide 14..25

FT /note= "p145 conformational B-cell epitope"

PN W09611944-A1.

PD 25-APR-1996.

PF 16-OCT-1995; 95WO-AU000681.

PR 14-OCT-1994; 94AU-00008851.

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.

PA (UYME ) UNIV MELBOURNE.

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.

PA (CSLC-) CSL LTD.

PI Cooper JA, Relf WA, Good MF, Saul AJ;

DR WPI; 1996-221939/22.

PT New chimeric peptide(s) including a conformational epitope - inserted  
 PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determin. of minimal epitope(s) or for mapping amphipathic helices.

PS Example 12; Fig 1C; 99pp; English.

XX The present peptide is a chimaeric peptide (CP) contg. the Streptococcal  
 CC M protein peptide p145 (Pruksakorn et al., J. Immunol. 149: 2729-2735  
 CC (1992)). The CP comprises a B-cell conformational epitope from within  
 CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on

CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
XX  
SQ Sequence 28 AA;  
Query Match 81.2%; Score 52; DB 2; Length 28;  
Best Local Similarity 78.6%; Pred. No. 0.24;  
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ASREAKKQVEKALE 14  
Db 10 ASREAKKQVEKVK 23  
RESULT 32  
AAR97412  
ID AAR97412 standard; peptide; 12 AA.  
XX  
AC AAR97412;  
XX  
DT 27-AUG-2003 (revised)  
DT 02-DEC-1996 (first entry)  
XX  
DE Streptococcal M protein peptide, p145, fragment J(1)6.  
XX  
KM Streptococcal; M protein; peptide; p145, chimeric; chimeric; B-cell;  
KM conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KM diagnosis.  
XX  
OS Streptococcus sp.  
XX  
PN WC9611944-A1.  
XX  
PD 25-APR-1996.  
XX  
PF 16-OCT-1995; 95WO-AU000681.  
XX  
PR 14-OCT-1994; 94AU-00008851.  
XX  
PA (COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UTME ) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX  
DR WPI; 1996-221939/22.  
XX  
PT New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 13; Page 37; 99pp; English.  
XX  
CC The present peptide is a fragment of the Streptococcal M protein peptide  
CC p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735 (1992)), used in the  
CC construction of a novel chimeric peptide (Cp). The Cp comprises a B-cell  
CC conformational epitope from within p145, inserted into a 2nd peptide,  
CC pref. an alpha-helical coil based on the GCN4 leucine zipper peptide  
CC (AAR97395). The 2nd peptide has a similar conformation, enabling the  
CC epitope to be presented in an immunologically active conformation. The Cp  
CC can be used in a novel detection/mapping process, e.g. to determine the  
CC min. epitope required to induce opsonic antibodies (Ab), and in vaccines

CC against gp. A Streptococci. Ab raised against the CP can be used for  
CC immunotherapy and diagnosis, while the CP can be used diagnostically to  
CC detect Ab. (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 12 AA;  
Query Match 79.7%; Score 51; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ASREAKKQVEK 11  
Db 2 ASREAKKQVEK 12  
RESULT 33  
ADG62867  
ID ADG62867 standard; peptide; 21 AA.  
XX  
AC ADG62867;  
XX  
DT 11-MAR-2004 (first entry)  
DT  
XX  
DE Streptococcus sp. M6 protein peptide epitope (240-260).  
XX  
KM Antigen conjugate; immune response; infection; C repeat region;  
KM cholera toxin B; CTB.  
XX  
OS Streptococcus sp.  
XX  
PN US6602507-B1.  
XX  
PD 05-AUG-2003.  
XX  
PF 06-JAN-1995; 95US-00369295.  
XX  
PR 25-MAR-1988; 88US-00173380.  
PR 27-FEB-1989; 89US-00315588.  
PR 19-JUN-1990; 90US-00540101.  
PR 03-MAR-1992; 92US-00845865.  
PR 28-MAY-1993; 93US-00068598.  
PR 22-FEB-1994; 94US-00200913.  
XX  
PA (YTRQ ) UNIV ROCKEFELLER.  
XX  
PI Fischetti VA;  
XX  
DR WPI; 2003-895270/82.  
XX  
PT New antigen conjugates comprising polypeptides from the conserved exposed  
PT region of group A streptococcal M protein, useful for preparing vaccines  
PT for controlling streptococcal infection a mammal.  
XX  
PS Claim 5; Col 13; 12pp; English.  
XX  
CC The invention relates an antigen conjugate comprising a linkable carrier  
CC covalently bound to a polypeptide consisting of five or more amino acid  
CC residues from the conserved exposed region of group A streptococcal M  
CC protein. The antigen conjugate elicits a protective immune response to  
CC streptococcal infection in a mammal when administered mucosally. The  
CC antigen conjugates and polypeptides are useful for preparing vaccines for  
CC controlling streptococcal infection or protecting against streptococcal  
CC infection in a mammal. The polypeptides were all within the C repeat  
CC region of the M protein of type 6 Streptococci (strain 471). The  
CC polypeptides were purified by high performance liquid chromatography.  
CC Highly purified cholera toxin B (CTB) was derivatised at primary amino  
CC groups with the heterobifunctional cross-linking agent N'-succinimidyl 3-  
CC (2-pyridyldithio) propionate (SPDP). The synthetic peptides were mixed  
CC with SPDP-derivatised CTB at a 1.5:1 ratio. An average of 1.06-1.27  
CC peptide molecules were covalently linked per CTB monomer. The present  
CC sequence is M6 protein peptide epitope from strain D471 of group A  
CC Streptococcus sp.

SO Sequence 21 AA;

Query Match 75.0%; Score 48; DB 7; Length 21;  
 Best Local Similarity 90.9%; Pred. No. 0.79;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEK 11  
 |||||  
 11 ASREAKKQVEK 21

Db

RESULT 34  
 AAW04351  
 ID AAW04351 standard; protein; 28 AA.  
 XX  
 AC AAW04351;  
 XX  
 DT 02-DEC-1996 (first entry)  
 XX  
 DE Chimaeric peptide (J5) contg. Streptococcal M protein peptide p145.  
 XX  
 KW Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
 KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
 KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
 KM diagnosis.  
 XX  
 OS Synthetic.  
 XX  
 FT Key Location/Qualifiers  
 FT Peptide 13..24  
 /note= "p145 conformational B-cell epitope"  
 XX  
 XX W09611944-A1.  
 XX  
 PD 25-APR-1996.  
 XX  
 PF 16-OCT-1995; 95WO-AU000681.  
 XX  
 PR 14-OCT-1994; 94AU-00008851.  
 XX  
 PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (UYME ) UNIV MELBOURNE.  
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
 PA (CSLC-) CSL LTD.  
 XX  
 PI Cooper JA, Relf WA, Good MF, Saul AJ;  
 XX  
 DR WPI; 1996-221939/22.  
 XX  
 PT New chimeric peptide(s) including a conformational epitope - inserted  
 PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
 XX  
 PS Example 12; Fig 1C; 99pp; English.  
 XX  
 CC The present peptide is a chimeric peptide (CP) contg. the Streptococcal  
 CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
 CC (1992)). The CP comprises a B-cell conformational epitope from within  
 CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
 CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar  
 CC conformation, enabling the epitope to be presented in an immunologically  
 CC active conformation. The CP can be used in a novel detection/mapping  
 CC process, e.g. to determine the min. epitope required to induce opsonic  
 CC antibodies (Ab), and in vaccines against sp. A Streptococci. Ab raised  
 CC against the CP can be used for immunotherapy and diagnosis, while the CP  
 CC can be used diagnostically to detect Ab. The reactivity of the present  
 CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
 CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
 XX  
 SQ Sequence 28 AA;

Query Match 74.2%; Score 47.5; DB 2; Length 28;  
 Best Local Similarity 76.5%; Pred. No. 1.3;  
 Matches 13; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ASREAKKQVE--KALE 14  
 |||||  
 11 ASREAKKQVEDKVKQLE 27

Db

RESULT 35  
 AAR97411  
 ID AAR97411 standard; peptide; 12 AA.  
 XX  
 AC AAR97411;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 02-DEC-1996 (first entry)  
 XX  
 DE Streptococcal M protein peptide, p145, fragment 5(1)5.  
 XX  
 KW Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
 KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
 KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
 KM diagnosis.  
 XX  
 OS Streptococcus sp.  
 XX  
 PN W09611944-A1.  
 XX  
 PD 25-APR-1996.  
 XX  
 PF 16-OCT-1995; 95WO-AU000681.  
 XX  
 PR 14-OCT-1994; 94AU-00008851.  
 XX  
 PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (UYME ) UNIV MELBOURNE.  
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
 PA (CSLC-) CSL LTD.  
 XX  
 PI Cooper JA, Relf WA, Good MF, Saul AJ;  
 XX  
 DR WPI; 1996-221939/22.  
 XX  
 PT New chimeric peptide(s) including a conformational epitope - inserted  
 PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
 XX  
 PS Example 13; Page 37; 99pp; English.  
 XX  
 CC The present peptide is a fragment of the Streptococcal M protein peptide  
 CC p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735 (1992)), used in the  
 CC construction of a novel chimaeric peptide (CP). The CP comprises a B-cell  
 CC conformational epitope from within p145, inserted into a 2nd peptide,  
 CC pref. an alpha-helical coil based on the GCN4 leucine zipper peptide  
 CC (AAR97395). The 2nd peptide has a similar conformation, enabling the  
 CC epitope to be presented in an immunologically active conformation. The CP  
 CC can be used in a novel detection/mapping process, e.g. to determine the  
 CC min. epitope required to induce opsonic antibodies (Ab), and in vaccines  
 CC against sp. A Streptococci. Ab raised against the CP can be used for  
 CC immunotherapy and diagnosis, while the CP can be used diagnostically to  
 CC detect Ab. (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 12 AA;

Query Match 71.9%; Score 46; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.92;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVE 10  
 |||||

Db 3 ASREAKKOVE 12

RESULT 36  
AAR97426  
ID AAR97426 standard; peptide: 20 AA.

XX AAR97426;  
AC  
XX  
XX 27-AUG-2003 (revised)  
DT 02-DEC-1996 (first entry)  
XX

DE Streptococcal M protein peptide 146, for use in chimeric peptide.  
XX  
XX Streptococcal; M protein; peptide; 146; chimeric; chimeric; B-cell;  
KM conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KM diagnosis.

OS Streptococcus sp.  
XX  
XX WO9611944-A1.  
XX  
XX 25-APR-1996.  
PD  
XX 16-OCT-1995; 95WO-AU000681.  
XX  
XX 14-OCT-1994; 94AU-00008851.  
XX  
XX (COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (COUNCIL COMMONWEALTH SCI & IND RES ORG.  
PA (TIME) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX  
XX WPI; 1996-221939/22.  
DR

PT New chimeric peptide(s) including a conformational epitope - inserted  
into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
XX  
XX Example 13; Page 38; 99pp; English.  
PS  
XX  
XX The present peptide is the Streptococcal M protein peptide 146 used in  
CC the construction of a novel chimeric peptide (CP). The CP comprises a B-  
CC cell conformational epitope from within the present peptide, inserted  
CC into a 2nd peptide, pref. an alpha-helical coil based on the GCN4 leucine  
CC zipper peptide (AAR97395). The 2nd peptide has a similar conformation,  
CC enabling the epitope to be presented in an immunologically active  
CC conformation. The CP can be used in a novel detection/mapping process,  
CC e.g. to determine the min. epitope required to induce opsonic antibodies  
CC (Ab), and in vaccines against gp. A Streptococci. Ab raised against the  
CC CP can be used for immunotherapy and diagnosis, while the CP can be used  
CC diagnostically to detect Ab. (Updated on 27-AUG-2003 to correct OS  
CC field.)  
XX  
XX

SO Sequence 20 AA;  
Query Match 71.9%; Score 46; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKKOVERALE 14  
Db 1 AKKOVERALE 10

RESULT 37  
ADN22689  
ID ADN22689 standard; protein: 546 AA.

XX  
AC ADN22689;  
XX  
XX  
XX 02-DEC-2004 (first entry)  
DT  
XX  
XX  
DE Bacterial polypeptide #5342.

XX Recombinant DNA construct; transformed plant; improved plant property;  
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KM pathogen tolerance; pest tolerance; plant disease resistance;  
KM cell cycle pathway modification; plant growth regulator;  
KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KM bacterial polypeptide.  
XX  
XX Bacteria.  
OS  
XX  
XX US2003233675-A1.  
PN  
XX 18-DEC-2003.  
PD  
XX 20-FEB-2003; 2003US-00369493.  
PF  
XX 21-FEB-2002; 2002US-0360039P.  
PR

XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SIAT/) SIATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
XX  
PI Cao Y, Hinkle GJ, Siater SC, Chen X, Goldman BS;  
XX  
XX WPI; 2004-061375/06.  
DR

PT New recombinant DNA construct comprising a promoter positioned to provide  
expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX  
XX Claim 1; SEQ ID NO 5342; 122pp; English.  
PS  
XX

CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
XX

SO Sequence 546 AA;  
Query Match 70.3%; Score 45; DB 8; Length 546;  
Best Local Similarity 75.0%; Pred. No. 75;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 REAKKOVERALE 14



XX 12-FEB-2002; 2002WO-IB002069.  
PF  
XX  
PR 12-FEB-2001; 2001GB-00003424.  
XX  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Fontana MR, Pizsa M, Maignani V, Monaci E;  
XX  
DR WPI; 2003-058415/05.  
XX  
DR N-PSDB; AB241921.  
XX  
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
XX  
PT medicament for treating or preventing N. gonorrhoeae infection.  
XX  
PS Disclosure; Page 802; 815pp; English.  
XX  
XX  
CC The present invention relates to proteins from Neisseria gonorrhoeae.  
CC Also disclosed are the nucleic acid molecules encoding the proteins and  
CC antibodies that specifically bind to the proteins. The composition  
CC comprising the protein, nucleic acid or antibody is useful for the  
CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
CC infection, this may be in the form of a vaccine or gene therapy.  
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid  
CC molecules of the invention  
XX  
SQ Sequence 153 AA;  
XX  
Query Match 68.8%; Score 44; DB 6; Length 153;  
Best Local Similarity 66.7%; Pred. No. 29;  
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SREAKQVERKL 13  
|||::|||:  
Db 42 SREARREYERKM 53

Search completed: June 13, 2005, 20:40:50  
Job time : 78.516 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 13, 2005, 20:37:05 ; Search time 19.5349 Seconds  
(without alignments)  
53.498 Million cell updates/sec

Title: US-10-706-275-1

Perfect score: 64

Sequence: 1 ASREAKKQVERALE 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	20	3	US-08-817-811-1
2	64	100.0	236	3	US-08-937-271-11
3	64	100.0	254	4	US-08-914-479A-4
4	64	100.0	284	4	US-08-914-479A-6
5	64	100.0	305	3	US-08-937-271-10
6	64	100.0	440	4	US-08-302-756E-35
7	64	100.0	443	2	US-08-795-475-6
8	64	100.0	443	4	US-08-325-278B-6
9	56	87.5	12	3	US-08-817-811-34
10	55	85.9	12	3	US-08-817-811-11
11	55	85.9	12	3	US-08-817-811-32
12	55	85.9	12	3	US-08-817-811-33
13	52	81.2	28	3	US-08-817-811-18
14	51	79.7	12	3	US-08-817-811-31
15	47.5	74.2	28	3	US-08-817-811-17
16	46	71.9	12	3	US-08-817-811-30
17	46	71.9	12	3	US-08-817-811-10
18	46	71.9	20	3	US-08-817-811-38
19	44.5	69.5	28	3	US-08-817-811-16
20	43	67.2	94	3	US-08-937-271-18
21	43	67.2	343	3	US-08-937-271-17
22	43	67.2	683	6	US-08-937-271-17
23	43	67.2	683	6	US-08-464-531-83
24	42	65.6	65	1	US-08-461-598-83
25	42	65.6	65	2	US-08-322-137-83
26	42	65.6	65	3	US-08-307-896-1
27	42	65.6	380	3	US-08-307-896-1

28	42	65.6	394	4	US-09-442-349A-106	Sequence 106, App
29	42	65.6	394	4	US-09-963-137-192	Sequence 192, App
30	42	65.6	394	4	US-09-963-137-194	Sequence 194, App
31	42	65.6	394	5	PCT-US95-11808-1	Sequence 1, Appli
32	42	65.6	409	4	US-09-902-540-12745	Sequence 12745, A
33	42	65.6	775	4	US-09-513-838-6	Sequence 6, Appli
34	42	65.6	1181	4	US-09-826-509-587	Sequence 587, App
35	41.5	64.8	28	3	US-08-817-811-29	Sequence 67, Appl
36	41	64.1	12	3	US-08-960-022-14	Sequence 14, Appl
37	41	64.1	514	2	US-08-817-811-15	Sequence 15, Appl
38	40	62.5	28	3	US-09-583-110-4222	Sequence 4222, Ap
39	39	60.9	64	4	US-09-513-999C-4696	Sequence 4696, Ap
40	39	60.9	85	4	US-09-107-433-3416	Sequence 3416, Ap
41	39	60.9	96	4	US-09-270-767-45471	Sequence 45471, A
42	39	60.9	459	4	US-08-817-811-8	Sequence 8, Appli
43	38	59.4	8	3	US-08-817-811-14	Sequence 8, Appli
44	38	59.4	19	6	5304631-14	Patent No. 5304631
45	38	59.4	19	6	5304631-14	Patent No. 5304631

## ALIGNMENTS

RESULT 1  
US-08-817-811-1  
; Sequence 1, Application US/08817811  
; Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-1

Query Match 100.0%; Score 64; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.00068;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
Db 7 ASREAKKQVEKALE 20

## RESULT 2

US-08-937-271-11  
Sequence 11, Application US/08937271  
Patent No. 6063386  
GENERAL INFORMATION:  
APPLICANT: Dale, James B.  
APPLICANT: Lederer, James W.  
TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN  
TITLE OF INVENTION: VACCINE  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,271  
FILING DATE: 15-SEP-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Stephen J.  
REGISTRATION NUMBER: 43,058  
REFERENCE/DOCKET NUMBER: 481112.405C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-937-271-11

Query Match 100.0%; Score 64; DB 3; Length 236;  
Best Local Similarity 100.0%; Pred. No. 0.0084;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
Db 87 ASREAKKQVEKALE 100

## RESULT 3

US-08-914-479A-4  
Sequence 4, Application US/08914479A  
Patent No. 6419932  
GENERAL INFORMATION:  
APPLICANT: Dale, James B.  
TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER  
TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE  
FILE REFERENCE: 481112.404C2  
CURRENT APPLICATION NUMBER: US/08/914,479A  
CURRENT FILING DATE: 1997-08-19  
PRIOR APPLICATION NUMBER: 08/409,270  
PRIOR FILING DATE: 1995-03-23  
PRIOR APPLICATION NUMBER: 07/945,860  
PRIOR FILING DATE: 1992-09-16  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4

LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An antigen of M5 and a carrier of the  
US-08-914-479A-4

Query Match 100.0%; Score 64; DB 4; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.009;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
Db 105 ASREAKKQVEKALE 118

## RESULT 4

US-08-914-479A-6  
Sequence 6, Application US/08914479A  
Patent No. 6419932  
GENERAL INFORMATION:  
APPLICANT: Dale, James B.  
TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER  
TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE  
FILE REFERENCE: 481112.404C2  
CURRENT APPLICATION NUMBER: US/08/914,479A  
CURRENT FILING DATE: 1997-08-19  
PRIOR APPLICATION NUMBER: 08/409,270  
PRIOR FILING DATE: 1995-03-23  
PRIOR APPLICATION NUMBER: 07/945,860  
PRIOR FILING DATE: 1992-09-16  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An antigen of three fragments of M5 and a carrier  
US-08-914-479A-6

Query Match 100.0%; Score 64; DB 4; Length 284;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
Db 135 ASREAKKQVEKALE 148

## RESULT 5

US-08-937-271-10  
Sequence 10, Application US/08937271  
Patent No. 6063386  
GENERAL INFORMATION:  
APPLICANT: Dale, James B.  
APPLICANT: Lederer, James W.  
TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN  
TITLE OF INVENTION: VACCINE  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,271  
FILING DATE: 15-SEP-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Rosenman, Stephen J.  
REGISTRATION NUMBER: 43,058  
REFERENCE/DOCKET NUMBER: 481112.405C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-937-271-10

Query Match 100.0%; Score 64; DB 3; Length 305;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASRAKKQVEKALE 14  
|||  
DB 156 ASRAKKQVEKALE 169

RESULT 6  
US-08-302-756E-35  
Sequence 35, Application US/08302756E  
Patent No. 6737521  
GENERAL INFORMATION:  
APPLICANT: FISCHETTI, Vincent A.  
APPLICANT: POZZI, Gianni  
TITLE OF INVENTION: DELIVERY AND EXPRESSION OF A HYBRID SURFACE PROTEIN ON  
FILE REFERENCE: 016921-076  
CURRENT APPLICATION NUMBER: US/08/302,756E  
CURRENT FILING DATE: 1995-03-07  
PRIOR APPLICATION NUMBER: US 07/522,440  
PRIOR FILING DATE: 1990-05-11  
PRIOR APPLICATION NUMBER: US 07/742,199  
PRIOR FILING DATE: 1991-08-05  
PRIOR APPLICATION NUMBER: US 07/814,823  
PRIOR FILING DATE: 1991-12-23  
PRIOR APPLICATION NUMBER: US 07/851,082  
PRIOR FILING DATE: 1992-03-13  
PRIOR APPLICATION NUMBER: PCT/US93/02355  
PRIOR FILING DATE: 1993-03-12  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 35  
LENGTH: 440  
TYPE: PRT  
ORGANISM: S. pyogenes  
US-08-302-756E-35

Query Match 100.0%; Score 64; DB 4; Length 440;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASRAKKQVEKALE 14  
|||  
DB 292 ASRAKKQVEKALE 305

RESULT 7  
US-08-795-475-6  
Sequence 6, Application US/08795475  
Patent No. 5965390

GENERAL INFORMATION:  
APPLICANT: Bjorck, Lars  
APPLICANT: Sjobring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,475  
FILING DATE: 11-FEB-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McMaisters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 100084.402D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 443 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-795-475-6

Query Match 100.0%; Score 64; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASRAKKQVEKALE 14  
|||  
DB 294 ASRAKKQVEKALE 307

RESULT 8  
US-08-325-278B-6  
Sequence 6, Application US/08325278B  
Patent No. 6822075  
GENERAL INFORMATION:  
APPLICANT: Bjorck, Lars  
APPLICANT: Sjobring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Fifth Avenue Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278B  
FILING DATE: 26-OCT-1994  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332

```

; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-325-278B-6

Query Match      100.0%; Score 64; DB 4; Length 443;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASREAKQVEKALE 14
Db      294 ASREAKQVEKALE 307

RESULT 9
US-08-817-811-34
; Sequence 34, Application US/08817811
; Patent No. 6174528
; GENERAL INFORMATION:
; APPLICANT: Cooper, Juan A.
; APPLICANT: Relf, Wendy A.
; APPLICANT: Good, Michael F.
; APPLICANT: Saul, Allan J.
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,811
; FILING DATE: 14-APR-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 96/11944
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: FBRC:005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-817-811-34

Query Match      87.5%; Score 56; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 REAKQVEKALE 14
```

```

Db      1 REAKQVEKALE 12

RESULT 10
US-08-817-811-11
; Sequence 11, Application US/08817811
; Patent No. 6174528
; GENERAL INFORMATION:
; APPLICANT: Cooper, Juan A.
; APPLICANT: Relf, Wendy A.
; APPLICANT: Good, Michael F.
; APPLICANT: Saul, Allan J.
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,811
; FILING DATE: 14-APR-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 96/11944
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: FBRC:005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-817-811-11

Query Match      85.9%; Score 55; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASREAKQVEKA 12
Db      1 ASREAKQVEKA 12

RESULT 11
US-08-817-811-32
; Sequence 32, Application US/08817811
; Patent No. 6174528
; GENERAL INFORMATION:
; APPLICANT: Cooper, Juan A.
; APPLICANT: Relf, Wendy A.
; APPLICANT: Good, Michael F.
; APPLICANT: Saul, Allan J.
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
```

STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-32

Query Match 85.9%; Score 55; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKA 12  
Db 1 ASREAKQVEKA 12

RESULT 12  
US-08-817-811-33  
Sequence 33, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-33

Query Match 85.9%; Score 55; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SREAKQVEKAL 13  
Db 1 SREAKQVEKAL 12

RESULT 13  
US-08-817-811-18  
Sequence 18, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-18

Query Match 81.2%; Score 52; DB 3; Length 28;  
Best Local Similarity 76.6%; Pred. No. 0.075;  
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASREAKKOVERKALE 14  
DB 10 ASREAKKOVERKVK 23

## RESULT 14

US-08-817-811-31  
; Sequence 31, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Relf, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,811  
; FILING DATE: 14-APR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 96/11944  
; FILING DATE: 25-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: FBRC:005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-817-811-31  
Query Match 79.7%; Score 51; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 15

US-08-817-811-17  
; Sequence 17, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Relf, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:

QY 1 ASREAKKOVER 11  
DB 2 ASREAKKOVER 12

ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-17  
Query Match 74.2%; Score 47.5; DB 3; Length 28;  
Best Local Similarity 76.5%; Pred. No. 0.38;  
Matches 13; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ASREAKKOVE--KALE 14  
DB 11 ASREAKKOVEDKVKOLE 27

## RESULT 16

US-08-817-811-10  
; Sequence 10, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Relf, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,811  
; FILING DATE: 14-APR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 96/11944  
; FILING DATE: 25-APR-1996

```
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: FIRC:005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-817-811-10

Query Match      71.9%; Score 46; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASREAKQVE 10
      |||||
Db      3 ASREAKQVE 12

RESULT 17
US-08-817-811-30
Sequence 30, Application US/08817811
Patent No. 6174528
GENERAL INFORMATION:
APPLICANT: Cooper, Juan A.
APPLICANT: Relf, Wendy A.
APPLICANT: Good, Michael F.
APPLICANT: Saul, Allan J.
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
NUMBER OF SEQUENCES: 97
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,811
FILING DATE: 14-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 96/11944
FILING DATE: 25-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: FIRC:005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-817-811-30

Query Match      71.9%; Score 46; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 ASREAKQVE 10
      |||||
Db      3 ASREAKQVE 12

RESULT 18
US-08-817-811-38
Sequence 38, Application US/08817811
Patent No. 6174528
GENERAL INFORMATION:
APPLICANT: Cooper, Juan A.
APPLICANT: Relf, Wendy A.
APPLICANT: Good, Michael F.
APPLICANT: Saul, Allan J.
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
NUMBER OF SEQUENCES: 97
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,811
FILING DATE: 14-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 96/11944
FILING DATE: 25-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: FIRC:005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-817-811-38

Query Match      71.9%; Score 46; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 AKKQVERALE 14
      |||||
Db      1 AKKQVERALE 10

RESULT 19
US-08-817-811-16
Sequence 16, Application US/08817811
Patent No. 6174528
GENERAL INFORMATION:
APPLICANT: Cooper, Juan A.
APPLICANT: Relf, Wendy A.
APPLICANT: Good, Michael F.
APPLICANT: Saul, Allan J.
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
NUMBER OF SEQUENCES: 97
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,811
FILING DATE: 14-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 96/11944
FILING DATE: 25-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: FBRC:005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
FAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-817-811-16

Query Match 69.5%; Score 44.5; DB 3; Length 28;
Best Local Similarity 70.6%; Pred. No. 1.1;
Matches 12; Conservative 1; Mismatches 1; Indels 3; Gaps 14

CY 1 ASREAKQVE--KALE 14
|||||: ||
Db 12 ASREAKQVQDKVKQLE 28

RESULT 20
US-08-937-271-18
Sequence 18, Application US/08937271
Patent No. 6063386
GENERAL INFORMATION:
APPLICANT: Dale, James B.
APPLICANT: Lederer, James W.
TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,271
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Rosenman, Stephen J.
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 481112.405C1

```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 94 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-937-271-18
Query Match 67.2%; Score 43; DB 3; Length 94;
Best Local Similarity 90.0%; Pred. No. 6.8;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVE 10
|||||:|
Db 17 ASREAKKQLE 26

RESULT 21
US-08-937-271-17
: Sequence 17, Application US/08937271
: Patent No. 6063386
: GENERAL INFORMATION:
: APPLICANT: Dale, James B.
: APPLICANT: Lederer, James W.
: TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN
: TITLE OF INVENTION: VACCINE
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED and BERRY
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/937,271
: FILING DATE: 15-SEP-1997
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Rosenman, Stephen J.
: REGISTRATION NUMBER: 43,058
: REFERENCE/DOCKET NUMBER: 481112.405C1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 343 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-937-271-17
Query Match 67.2%; Score 43; DB 3; Length 343;
Best Local Similarity 90.0%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVE 10
|||||:|
Db 266 ASREAKKQLE 275

RESULT 22
5210183-3
Patent No. 5210183

```



APPLICANT: LINDAHL, GUNNAR;FRITZ, ELISABET;HEDEN, LARS-OLOF  
TITLE OF INVENTION: PROTEIN ARP, WITH IMMUNOGLOBULIN A  
BINDING ACTIVITY, THE CORRESPONDING VECTORS AND HOSTS, REAGENT  
KIT AND PHARMACEUTICAL COMPOSITION  
NUMBER OF SEQUENCES: 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/270,099  
FILING DATE: 14-NOV-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 186,097  
FILING DATE: 25-APR-1988  
SEQ ID NO:3:  
LENGTH: 683

Query Match 67.2%; Score 43; DB 6; Length 683;  
Best Local Similarity 76.9%; Pred. No. 51;  
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKAL 13  
DB 515 ASREAKKQVEADL 527

RESULT 23  
5210183-3  
PATENT NO. 5210183  
APPLICANT: LINDAHL, GUNNAR;FRITZ, ELISABET;HEDEN, LARS-OLOF  
TITLE OF INVENTION: PROTEIN ARP, WITH IMMUNOGLOBULIN A  
BINDING ACTIVITY, THE CORRESPONDING VECTORS AND HOSTS, REAGENT  
KIT AND PHARMACEUTICAL COMPOSITION  
NUMBER OF SEQUENCES: 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/270,099  
FILING DATE: 14-NOV-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 186,097  
FILING DATE: 25-APR-1988  
SEQ ID NO:3:  
LENGTH: 683

Query Match 67.2%; Score 43; DB 6; Length 683;  
Best Local Similarity 76.9%; Pred. No. 51;  
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKAL 13  
DB 515 ASREAKKQVEADL 527

RESULT 24  
US-08-464-531-83  
SEQUENCE 83, APPLICATION US/08464531  
PATENT NO. 5789184  
GENERAL INFORMATION:  
APPLICANT: FOWLES, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: TRUSHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,531  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/322,137  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: FOWLES-2G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 65 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-531-83

Query Match 65.6%; Score 42; DB 1; Length 65;  
Best Local Similarity 57.1%; Pred. No. 6.7;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKALE 14  
DB 18 AGRANKKIKKQLO 31

RESULT 25  
US-08-461-598-83  
SEQUENCE 83, APPLICATION US/08461598  
PATENT NO. 5876951  
GENERAL INFORMATION:  
APPLICANT: FOWLES, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: TRUSHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,598
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/322,137
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,313
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,328
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/041,431
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: FOLWKS=2F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-461-598-83

Query Match      65.6%; Score 42; DB 2; Length 65;
Best Local Similarity 57.1%; Pred. No. 6.7;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
QY      1 ASREAKKQVEKALE 14
      |||::|||:
Db      18 AGRANKKIEKQLO 31
```

```
RESULT 26
US-08-322-137-83
Sequence 83, Application US/08322137
Patent No. 6100042
GENERAL INFORMATION:
APPLICANT: FOWLES, Dana M.
APPLICANT: BROACH, Jim
APPLICANT: MANFREDI, John
APPLICANT: KLEIN, Christine
APPLICANT: MURPHY, Andrew J.
APPLICANT: PAUL, Jeremy
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
TITLE OF INVENTION: PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,137
FILING DATE: 13-OCT-1994
```

```
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,313
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,328
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/041,431
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: FOLWKS=2C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-322-137-83

Query Match      65.6%; Score 42; DB 3; Length 65;
Best Local Similarity 57.1%; Pred. No. 6.7;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
QY      1 ASREAKKQVEKALE 14
      |||::|||:
Db      18 AGRANKKIEKQLO 31
```

```
RESULT 27
US-08-307-896-1
Sequence 1, Application US/08307896C
Patent No. 6034071
GENERAL INFORMATION:
APPLICANT: Iyengar, Srinivas Ravi
TITLE OF INVENTION: MOTANT ACTIVATED GSALPHA AND ADENYLYL
FILE REFERENCE: 29770
CURRENT APPLICATION NUMBER: US/08/307,896C
CURRENT FILING DATE: 1994-09-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-08-307-896-1
```

```
Query Match      65.6%; Score 42; DB 3; Length 380;
Best Local Similarity 57.1%; Pred. No. 40;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
QY      1 ASREAKKQVEKALE 14
      |||::|||:
Db      18 AGRANKKIEKQLO 31
```

```
RESULT 28
US-09-442-349A-106
Sequence 106, Application US/09442349A
Patent No. 6462178
GENERAL INFORMATION:
APPLICANT: Wong, Yung H
TITLE OF INVENTION: G Protein
FILE REFERENCE: M99/0101/US
CURRENT APPLICATION NUMBER: US/09/442,349A
```



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: GALPRA
PCT-US95-11808-1

Query Match          65.6%; Score 42; DB 5; Length 394;
Best Local Similarity 57.1%; Pred. No. 42;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 ASREAKKQVEKALE 14
Db      18 AGRANKKIKKQIQ 31

RESULT 32
US-09-902-540-12745
; Sequence 12745, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12745
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12745

Query Match          65.6%; Score 42; DB 4; Length 409;
Best Local Similarity 57.1%; Pred. No. 43;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 ASREAKKQVEKALE 14
Db      230 AEGPARQVERELE 243

RESULT 33
US-09-513-838-6
; Sequence 6, Application US/09513838
; Patent No. 6420563
; GENERAL INFORMATION:
; APPLICANT: Beeley, Nigel R
; APPLICANT: Behan, Dominic P
; APPLICANT: Chalmers, Derek T
; APPLICANT: Menzaghi, Frederique
; APPLICANT: Strah-Pleyvet, Sonja
; TITLE OF INVENTION: Small Molecule Modulators Of G Protein-Coupled Receptor
; FILE REFERENCE: AREN0058
; CURRENT APPLICATION NUMBER: US/09/513,838
; PRIOR FILING DATE: 2000-02-25
; EARLIER APPLICATION NUMBER: 09/364,425
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/094,879
; EARLIER FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: 60/106,300
; EARLIER FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/110,906
; EARLIER FILING DATE: 1998-12-04
```

```

; EARLIER APPLICATION NUMBER: 60/121,851
; EARLIER FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 60/173,850
; EARLIER FILING DATE: 1999-12-30
; EARLIER APPLICATION NUMBER: 60/174,428
; EARLIER FILING DATE: 2000-01-04
; EARLIER APPLICATION NUMBER: 09/364,425
; EARLIER FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-838-6

Query Match          65.6%; Score 42; DB 4; Length 775;
Best Local Similarity 57.1%; Pred. No. 83;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 ASREAKKQVEKALE 14
Db      399 AGRANKKIKKQIQ 412

RESULT 34
US-09-826-509-587
; Sequence 587, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Brulsma, Karin
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 587
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-587

Query Match          65.6%; Score 42; DB 4; Length 1181;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 ASREAKKQVEKALE 14
Db      805 AGRANKKIKKQIQ 818

RESULT 35
US-08-817-811-67
; Sequence 67, Application US/08817811
; Patent No. 6174528
; GENERAL INFORMATION:
; APPLICANT: Cooper, Juan A.
; APPLICANT: Relf, Wendy A.
; APPLICANT: Good, Michael F.
; APPLICANT: Saul, Allan J.
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
; TITLE OF INVENTION: COMPETING SAME
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
```

STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FIRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-67

Query Match 64.8%; Score 41.5; DB 3; Length 28;  
Best Local Similarity 64.7%; Pred. No. 3.4;  
Matches 11; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

OY 1 ASREAKQVE--KALE 14  
|||  
Db 12 ASREAKQLQDKVKOLE 28

RESULT 36  
US-08-817-811-29  
Sequence 29, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Kelt, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FIRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-29

Query Match 64.1%; Score 41; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASREAKQV 9  
|||  
Db 4 ASREAKQV 12

RESULT 37  
US-08-960-022-14  
Sequence 14, Application US/08960022  
Patent No. 5976837  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John M.  
APPLICANT: Lavallee, Edward R.  
APPLICANT: Racie, Lisa A.  
APPLICANT: Merberg, David  
APPLICANT: Treacy, Maurice  
APPLICANT: Spaulding, Vikki  
APPLICANT: Agostino, Michael J.  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/960.022  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-960-022-14

Query Match 64.1%; Score 41; DB 2; Length 514;  
Best Local Similarity 66.7%; Pred. No. 79;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 REAKKOVEKALE 14  
 |||:|:|:  
 Db 407 REAKEQERAME 418

RESULT 38  
 US-08-817-811-15  
 ; Sequence 15, Application US/08817811  
 ; Patent No. 6174528  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cooper, Juan A.  
 ; APPLICANT: Relif, Wendy A.  
 ; APPLICANT: Good, Michael F.  
 ; APPLICANT: Saul, Allan J.  
 ; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
 ; TITLE OF INVENTION: COMPRISING SAME  
 ; NUMBER OF SEQUENCES: 97  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Arnold, White & Durkee  
 ; STREET: P.O. Box 4433  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77210  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/817,811  
 ; FILING DATE: 14-APR-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO 96/11944  
 ; FILING DATE: 25-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Highlander, Steven L.  
 ; REGISTRATION NUMBER: 37,642  
 ; REFERENCE/DOCKET NUMBER: PARC:005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 512/418-3000  
 ; TELEFAX: 512/474-7577  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 28 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; US-08-817-811-15

Query Match 62.5%; Score 40; DB 3; Length 28;  
 Best Local Similarity 80.0%; Pred. No. 5.9;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVE 10  
 |||:|:|:  
 Db 13 ASREAKKQLO 22

RESULT 39  
 US-09-583-110-4222  
 ; Sequence 4222, Application US/09583110  
 ; Patent No. 6699703  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al.  
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
 ; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
 ; FILE REFERENCE: PAT00-07A  
 ; CURRENT APPLICATION NUMBER: US/09/583,110  
 ; CURRENT FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30  
 ; PRIOR APPLICATION NUMBER: US 60/085,131  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: US 60/051,553  
 ; PRIOR FILING DATE: 1997-07-02  
 ; NUMBER OF SEQ ID NOS: 5322  
 ; SEQ ID NO 4222  
 ; LENGTH: 64  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 ; US-09-583-110-4222

Query Match 60.9%; Score 39; DB 4; Length 64;  
 Best Local Similarity 69.2%; Pred. No. 20;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASREAKKQVEKAL 13  
 |||:|:|:  
 Db 5 ASREAKKQVSKAL 17

RESULT 40  
 US-09-513-999C-4696  
 ; Sequence 4696, Application US/09513999C  
 ; Patent No. 6783961  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J. B.  
 ; APPLICANT: Duclerc, A.  
 ; APPLICANT: Giordano, J. Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6783961  
 ; FILE REFERENCE: 59.US2.RBG  
 ; CURRENT APPLICATION NUMBER: US/09/513,999C  
 ; CURRENT FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/122,487  
 ; PRIOR FILING DATE: 1999-02-26  
 ; NUMBER OF SEQ ID NOS: 36681  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 4696  
 ; LENGTH: 85  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIGNAL  
 ; LOCATION: -45...-1  
 ; OTHER INFORMATION: score 6.7  
 ; OTHER INFORMATION: seq PMLGLAFAFWIWS/RE  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: 35  
 ; OTHER INFORMATION: Xaa=Asp or Glu  
 ; US-09-513-999C-4696

Query Match 60.9%; Score 39; DB 4; Length 85;  
 Best Local Similarity 61.5%; Pred. No. 26;  
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SREAKKQVEKALE 14  
 |||:|:|:  
 Db 45 SRESQKEVEKERE 57

Search completed: June 13, 2005, 20:58:36  
 Job time : 20.5349 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:52:55 ; Search time 68.3721 Seconds  
(without alignments)  
78.492 Million cell updates/sec

Title: US-10-706-275-1

Perfect score: 64

Sequence: 1 ASREAXKQVEKALE 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38333425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppa/PC7\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppa/PC7US\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubppa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubppa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubppa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	14	17	US-10-706-275-1
2	64	100.0	20	13	US-10-044-034-22
3	64	100.0	20	17	US-10-706-275-5
4	64	100.0	29	17	US-10-706-275-2
5	64	100.0	29	17	US-10-706-275-15
6	64	100.0	254	13	US-10-141-627-4
7	64	100.0	284	13	US-10-141-627-6
8	64	100.0	443	8	US-08-325-278-6
9	64	100.0	553	16	US-10-474-792-672
10	64	100.0	558	17	US-10-732-923-3295
11	57	89.1	28	17	US-10-706-275-12

12	56	87.5	28	17	US-10-706-275-13	Sequence 13, Appl
13	56	87.5	28	17	US-10-706-275-14	Sequence 14, Appl
14	52	81.2	28	17	US-10-706-275-11	Sequence 11, Appl
15	47.5	74.2	28	17	US-10-706-275-10	Sequence 10, Appl
16	45	70.3	546	15	US-10-369-493-5342	Sequence 5342, Ap
17	44.5	69.5	28	17	US-10-706-275-9	Sequence 9, Appl1
18	42	65.6	65	15	US-09-309-196-83	Sequence 83, Appl
19	42	65.6	65	15	US-10-263-341-83	Sequence 83, Appl
20	42	65.6	65	16	US-10-600-003-83	Sequence 83, Appl
21	42	65.6	86	15	US-10-424-599-223904	Sequence 223904,
22	42	65.6	377	17	US-10-732-923-7618	Sequence 7618, Ap
23	42	65.6	379	10	US-09-952-680A-15	Sequence 15, Appl
24	42	65.6	379	16	US-10-408-765A-105	Sequence 105, App
25	42	65.6	379	16	US-10-215-982-15	Sequence 15, Appl
26	42	65.6	379	17	US-10-732-923-7963	Sequence 7963, Ap
27	42	65.6	379	17	US-10-732-923-8015	Sequence 8015, Ap
28	42	65.6	380	16	US-09-952-680A-16	Sequence 16, Appl
29	42	65.6	380	10	US-10-215-982-16	Sequence 16, Appl
30	42	65.6	380	17	US-10-732-923-8016	Sequence 8016, Ap
31	42	65.6	380	17	US-10-732-923-8055	Sequence 8055, Ap
32	42	65.6	384	17	US-10-732-923-8022	Sequence 8022, Ap
33	42	65.6	388	17	US-10-732-923-8054	Sequence 8054, Ap
34	42	65.6	394	10	US-09-952-680A-13	Sequence 13, Appl
35	42	65.6	394	10	US-09-963-131-192	Sequence 192, App
36	42	65.6	394	10	US-09-963-131-194	Sequence 194, App
37	42	65.6	394	15	US-10-116-275-187	Sequence 187, App
38	42	65.6	394	15	US-10-352-943-18	Sequence 18, Appl
39	42	65.6	394	16	US-10-215-982-13	Sequence 13, Appl
40	42	65.6	394	16	US-10-684-422-194	Sequence 194, App
41	42	65.6	394	17	US-10-732-923-7589	Sequence 7589, Ap
42	42	65.6	394	17	US-10-732-923-7592	Sequence 7592, Ap
43	42	65.6	394	17	US-10-732-923-7620	Sequence 7620, Ap
44	42	65.6	394	17	US-10-732-923-7653	Sequence 7653, Ap
45	42	65.6	394	17	US-10-732-923-8017	Sequence 8017, Ap

#### ALIGNMENTS

RESULT 1  
US-10-706-275-1  
Sequence 1, Application US/10706275  
Publication No. US20050002956A1  
GENERAL INFORMATION:  
APPLICANT: ID Biomedical Corporation of Quebec  
APPLICANT: The Council of the Queensland Institute of Medical Research  
APPLICANT: Lowell, George H.  
APPLICANT: Butte, David S.  
APPLICANT: White, Gregory L.  
APPLICANT: Good, Michael F.  
APPLICANT: Batzloff, Michael R.  
APPLICANT: Leanderson, Tomas B.  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 021989-000710US  
CURRENT APPLICATION NUMBER: US/10/706,275  
CURRENT FILING DATE: 2003-11-13  
PRIOR APPLICATION NUMBER: US 60/426,409  
PRIOR FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: AU 2002302132  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of S. pyogenes  
US-10-706-275-1  
Query Match 100.0%; Score 64; DB 17; Length 14;  
Best local Similarity 100.0%; Pred. No. 0.0024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVEKALE 14  
 |||||  
 Db 1 ASREAKKOVEKALE 14

# RESULT 2

US-10-044-034-22  
 ; Sequence 22, Application US/10044034  
 ; Publication No. US20020169264A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JACKSON, DAVID C.  
 ; APPLICANT: O'BRIEN-SIMPSON, NEIL M.  
 ; APPLICANT: BROWN, LORENA E.  
 ; APPLICANT: EDE, NICHOLAS J.  
 ; APPLICANT: BRANDT, EVELYN R.  
 ; APPLICANT: GOOD, MICHAEL F.  
 ; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES  
 ; FILE REFERENCE: FBRC:006  
 ; CURRENT APPLICATION NUMBER: US/10/044,034  
 ; CURRENT FILING DATE: 2002-01-11  
 ; PRIOR APPLICATION NUMBER: P05071  
 ; PRIOR FILING DATE: 1997-02-11  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 22  
 ; LENGTH: 20  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: Peptides  
 ; US-10-044-034-22

Query Match 100.0%; Score 64; DB 13; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0034;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVEKALE 14  
 |||||  
 Db 7 ASREAKKOVEKALE 20

# RESULT 3

US-10-706-275-5  
 ; Sequence 5, Application US/10706275  
 ; Publication No. US20050002956A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ID Biomedical Corporation of Quebec  
 ; APPLICANT: The Council of the Queensland Institute of Medical Research  
 ; APPLICANT: Lowell, George H.  
 ; APPLICANT: Burt, David S.  
 ; APPLICANT: White, Gregory L.  
 ; APPLICANT: Batzloff, Michael R.  
 ; APPLICANT: Leanderson, Tomas B.  
 ; TITLE OF INVENTION: Vaccine  
 ; FILE REFERENCE: 021989-000710US  
 ; CURRENT APPLICATION NUMBER: US/10/706,275  
 ; CURRENT FILING DATE: 2003-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/426,409  
 ; PRIOR FILING DATE: 2002-11-15  
 ; PRIOR APPLICATION NUMBER: AU 2002302132  
 ; PRIOR FILING DATE: 2002-11-15  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 20  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: antigenic peptide sequence p145  
 ; US-10-706-275-5

Query Match 100.0%; Score 64; DB 17; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0034;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVEKALE 14  
 |||||  
 Db 7 ASREAKKOVEKALE 20

# RESULT 4

US-10-706-275-2  
 ; Sequence 2, Application US/10706275  
 ; Publication No. US20050002956A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ID Biomedical Corporation of Quebec  
 ; APPLICANT: The Council of the Queensland Institute of Medical Research  
 ; APPLICANT: Lowell, George H.  
 ; APPLICANT: Burt, David S.  
 ; APPLICANT: White, Gregory L.  
 ; APPLICANT: Good, Michael F.  
 ; APPLICANT: Batzloff, Michael R.  
 ; APPLICANT: Leanderson, Tomas B.  
 ; TITLE OF INVENTION: Vaccine  
 ; FILE REFERENCE: 021989-000710US  
 ; CURRENT APPLICATION NUMBER: US/10/706,275  
 ; CURRENT FILING DATE: 2003-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/426,409  
 ; PRIOR FILING DATE: 2002-11-15  
 ; PRIOR APPLICATION NUMBER: AU 2002302132  
 ; PRIOR FILING DATE: 2002-11-15  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 29  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: antigenic peptide derivative of S. pyogenes with flanking sequen  
 ; OTHER INFORMATION: es  
 ; US-10-706-275-2

Query Match 100.0%; Score 64; DB 17; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVEKALE 14  
 |||||  
 Db 9 ASREAKKOVEKALE 22

# RESULT 5

US-10-706-275-15  
 ; Sequence 15, Application US/10706275  
 ; Publication No. US20050002956A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ID Biomedical Corporation of Quebec  
 ; APPLICANT: The Council of the Queensland Institute of Medical Research  
 ; APPLICANT: Lowell, George H.  
 ; APPLICANT: Burt, David S.  
 ; APPLICANT: White, Gregory L.  
 ; APPLICANT: Batzloff, Michael R.  
 ; APPLICANT: Leanderson, Tomas B.  
 ; TITLE OF INVENTION: Vaccine  
 ; FILE REFERENCE: 021989-000710US  
 ; CURRENT APPLICATION NUMBER: US/10/706,275  
 ; CURRENT FILING DATE: 2003-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/426,409  
 ; PRIOR FILING DATE: 2002-11-15  
 ; PRIOR APPLICATION NUMBER: AU 2002302132  
 ; PRIOR FILING DATE: 2002-11-15  
 ; NUMBER OF SEQ ID NOS: 15



SOFTWARE: Patentin version 3.1  
SEQ ID NO 15  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-15

Query Match 100.0%; Score 64; DB 17; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.005;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 9 ASREAKKQVEKALE 22

RESULT 6  
US-10-141-627-4  
Sequence 6, Application US/10141627  
Publication No. US20020176863A1  
GENERAL INFORMATION:  
APPLICANT: Dale, James B.  
TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER  
FILE REFERENCE: 481112.404C3  
CURRENT APPLICATION NUMBER: US/10/141.627  
CURRENT FILING DATE: 2002-05-07  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An antigen of M5 and a carrier of the  
OTHER INFORMATION: COOH-terminal portion of M5  
US-10-141-627-4

Query Match 100.0%; Score 64; DB 13; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.048;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 105 ASREAKKQVEKALE 118

RESULT 7  
US-10-141-627-6  
Sequence 6, Application US/10141627  
Publication No. US20020176863A1  
GENERAL INFORMATION:  
APPLICANT: Dale, James B.  
TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER  
FILE REFERENCE: 481112.404C3  
CURRENT APPLICATION NUMBER: US/10/141.627  
CURRENT FILING DATE: 2002-05-07  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An antigen of three fragments of M5 and a carrier  
OTHER INFORMATION: of the COOH-terminal portion of M5  
US-10-141-627-6

Query Match 100.0%; Score 64; DB 13; Length 284;  
Best Local Similarity 100.0%; Pred. No. 0.054;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ASREAKKQVEKALE 14  
Db 135 ASREAKKQVEKALE 148

RESULT 8  
US-08-325-278-6  
Sequence 6, Application US/08325278  
Publication No. US20030027283A1  
GENERAL INFORMATION:  
APPLICANT: Bjvick, Lars  
APPLICANT: Sjvdring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325.278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McMaisters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 450023.401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 443 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-325-278-6

Query Match 100.0%; Score 64; DB 8; Length 443;  
Best Local Similarity 100.0%; Pred. No. 0.085;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 294 ASREAKKQVEKALE 307

RESULT 9  
US-10-474-792-672  
Sequence 672, Application US/10474792  
Publication No. US20040236072A1  
GENERAL INFORMATION:  
APPLICANT: Olmsted, Stephen  
APPLICANT: Zagureky, Robert  
APPLICANT: Nickbarg, Elliot  
APPLICANT: Winter, Louie  
TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES  
FILE REFERENCE: AM 100399  
CURRENT APPLICATION NUMBER: US/10/474.792  
CURRENT FILING DATE: 2003-10-14  
NUMBER OF SEQ ID NOS: 674  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 672  
LENGTH: 553

```
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
; US-10-474-792-672
```

```
Query Match          100.0%; Score 64; DB 16; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ASREAKKQVEKALE 14
      |||||
Db      404 ASREAKKQVEKALE 417
```

```
RESULT 10
US-10-732-923-3295
; Sequence 3295, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3295
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-732-923-3295
```

```
Query Match          100.0%; Score 64; DB 17; Length 558;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ASREAKKQVEKALE 14
      |||||
Db      409 ASREAKKQVEKALE 422
```

```
RESULT 11
US-10-706-275-12
; Sequence 12, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Good, Michael F.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: AU 2002302132
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antigenic peptide sequence derivative of p145 base
US-10-706-275-12
Query Match          89.1%; Score 57; DB 17; Length 28;
```

```
Best Local Similarity 85.7%; Pred. No. 0.059;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ASREAKKQVEKALE 14
      |||||
Db      9 ASREAKKQVEKAVK 22
```

```
RESULT 12
US-10-706-275-13
; Sequence 13, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Good, Michael F.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: AU 2002302132
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antigenic peptide sequence derivative of p145 base
US-10-706-275-13
```

```
Query Match          87.5%; Score 56; DB 17; Length 28;
Best Local Similarity 92.3%; Pred. No. 0.085;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 SREAKKQVEKALE 14
      |||||
Db      9 SREAKKQVEKALK 21
```

```
RESULT 13
US-10-706-275-14
; Sequence 14, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Good, Michael F.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: AU 2002302132
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 28
```

```
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: antigenic peptide sequence derivative of p145 base
US-10-706-275-14
```

```
Query Match      87.5%; Score 56; DB 17; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 REAKKQVEKALE 14
DB      9 REAKKQVEKALE 20
```

```
RESULT 14
US-10-706-275-11
/ Sequence 11, Application US/10706275
/ Publication No. US2005002956A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: ID Biomedical Corporation of Quebec
/ APPLICANT: The Council of the Queensland Institute of Medical Research
/ APPLICANT: Lowell, George H.
/ APPLICANT: Burt, David S.
/ APPLICANT: White, Gregory L.
/ APPLICANT: Good, Michael F.
/ APPLICANT: Batzloff, Michael R.
/ APPLICANT: Leanderson, Tomas B.
/ TITLE OF INVENTION: Vaccine
/ FILE REFERENCE: 021989-000710US
/ CURRENT APPLICATION NUMBER: US/10/706,275
/ CURRENT FILING DATE: 2003-11-13
/ PRIOR APPLICATION NUMBER: US 60/426,409
/ PRIOR FILING DATE: 2002-11-15
/ PRIOR APPLICATION NUMBER: AU 2002302132
/ PRIOR FILING DATE: 2002-11-15
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11
/ LENGTH: 28
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: antigenic peptide sequence derivative of p145 base
US-10-706-275-11
```

```
Query Match      81.2%; Score 52; DB 17; Length 28;
Best Local Similarity 78.6%; Pred. No. 0.36;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 ASREAKKQVEKALE 14
DB     10 ASREAKKQVEKAKV 23
```

```
RESULT 15
US-10-706-275-10
/ Sequence 10, Application US/10706275
/ Publication No. US20050002956A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: ID Biomedical Corporation of Quebec
/ APPLICANT: The Council of the Queensland Institute of Medical Research
/ APPLICANT: Lowell, George H.
/ APPLICANT: Burt, David S.
/ APPLICANT: White, Gregory L.
/ APPLICANT: Good, Michael F.
/ APPLICANT: Batzloff, Michael R.
/ APPLICANT: Leanderson, Tomas B.
/ TITLE OF INVENTION: Vaccine
/ FILE REFERENCE: 021989-000710US
/ CURRENT APPLICATION NUMBER: US/10/706,275
/ CURRENT FILING DATE: 2003-11-13
/ PRIOR APPLICATION NUMBER: US 60/426,409
```

```
/ PRIOR FILING DATE: 2002-11-15
/ PRIOR APPLICATION NUMBER: AU 2002302132
/ PRIOR FILING DATE: 2002-11-15
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 10
/ LENGTH: 28
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: antigenic peptide sequence derivative of p145 base
US-10-706-275-10
```

```
Query Match      74.2%; Score 47.5; DB 17; Length 28;
Best Local Similarity 76.5%; Pred. No. 1.8;
Matches 13; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
```

```
OY      1 ASREAKKQVE--KALE 14
DB     11 ASREAKKQVEDKVKOLE 27
```

```
RESULT 16
US-10-369-493-5342
/ Sequence 5342, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 5342
/ LENGTH: 546
/ TYPE: PRT
/ ORGANISM: Caenorhabditis elegans
US-10-369-493-5342
```

```
Query Match      70.3%; Score 45; DB 15; Length 546;
Best Local Similarity 75.0%; Pred. No. 96;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      3 REAKKQVEKALE 14
DB     406 RERKQTEKELE 417
```

```
RESULT 17
US-10-706-275-9
/ Sequence 9, Application US/10706275
/ Publication No. US20050002956A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: ID Biomedical Corporation of Quebec
/ APPLICANT: The Council of the Queensland Institute of Medical Research
/ APPLICANT: Lowell, George H.
/ APPLICANT: Burt, David S.
/ APPLICANT: White, Gregory L.
/ APPLICANT: Good, Michael F.
/ APPLICANT: Batzloff, Michael R.
/ APPLICANT: Leanderson, Tomas B.
/ TITLE OF INVENTION: Vaccine
/ FILE REFERENCE: 021989-000710US
/ CURRENT APPLICATION NUMBER: US/10/706,275
/ CURRENT FILING DATE: 2003-11-13
/ PRIOR APPLICATION NUMBER: US 60/426,409
```

;; PRIOR FILING DATE: 2002-11-15  
;; PRIOR APPLICATION NUMBER: AU 2002302132  
;; PRIOR FILING DATE: 2002-11-15  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 9  
;; LENGTH: 28  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-9

Query Match 69.5%; Score 44.5; DB 17; Length 28;  
Best Local Similarity 70.6%; Pred. No. 5.2;  
Matches 12; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

Qy 1 ASREAKKQVE---KALE 14  
Db 12 ASREAKKQVDKVKQLE 28

RESULT 18  
US-09-309-196-83  
;; Sequence 83, Application US/09309196  
;; Publication No. US2003008380A1  
;; GENERAL INFORMATION:  
;; APPLICANT: FOWLES, Dana M.  
;; APPLICANT: BROACH, Jim  
;; APPLICANT: MANFREDI, John  
;; APPLICANT: KLEIN, Christine  
;; APPLICANT: MURPHY, Andrew J.  
;; APPLICANT: PAUL, Jeremy  
;; APPLICANT: TRUEHEART, Joshua  
;; TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
;; TITLE OF INVENTION: PHERMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
;; NUMBER OF SEQUENCES: 119  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: BROWDY AND NEIMARK  
;; STREET: 419 Seventh Street, N.W., Suite 300  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20004  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/309,196  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/322,137  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/190,328  
;; FILING DATE: 31-JAN-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/041,431  
;; FILING DATE: 31-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: COOPER, Iver P.  
;; REGISTRATION NUMBER: 28,005  
;; REFERENCE/DOCKET NUMBER: FOLWRES-2C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-628-5197  
;; TELEFAX: 202-737-3528  
;; TELEEX: 248633  
;; INFORMATION FOR SEQ ID NO: 83:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 65 amino acids

;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-309-196-83

Query Match 65.6%; Score 42; DB 10; Length 65;  
Best Local Similarity 57.1%; Pred. No. 31;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 18 AQREANKKIEKQLE 31

RESULT 19  
US-10-263-341-83  
;; Sequence 83, Application US/10263341  
;; Publication No. US20030203417A1  
;; GENERAL INFORMATION:  
;; APPLICANT: FOWLES, Dana M.  
;; APPLICANT: BROACH, Jim  
;; APPLICANT: MANFREDI, John  
;; APPLICANT: KLEIN, Christine  
;; APPLICANT: MURPHY, Andrew J.  
;; APPLICANT: PAUL, Jeremy  
;; APPLICANT: TRUEHEART, Joshua  
;; TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
;; TITLE OF INVENTION: PHERMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
;; NUMBER OF SEQUENCES: 119  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: BROWDY AND NEIMARK  
;; STREET: 419 Seventh Street, N.W., Suite 300  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20004  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/263,341  
;; FILING DATE: 01-Oct-2002  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/322,137  
;; FILING DATE: 13-OCT-1994  
;; APPLICATION NUMBER: US 08/309,313  
;; FILING DATE: 20-SEP-1994  
;; APPLICATION NUMBER: US 08/190,328  
;; FILING DATE: 31-JAN-1994  
;; APPLICATION NUMBER: US 08/041,431  
;; FILING DATE: 31-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: COOPER, Iver P.  
;; REGISTRATION NUMBER: 28,005  
;; REFERENCE/DOCKET NUMBER: FOLWRES-2C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-628-5197  
;; TELEFAX: 202-737-3528  
;; TELEEX: 248633  
;; INFORMATION FOR SEQ ID NO: 83:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 65 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 83:  
US-10-263-341-83

Query Match 65.6%; Score 42; DB 15; Length 65;  
Best Local Similarity 57.1%; Pred. No. 31;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ASREAKKQVERKALE 14  
Db 18 AQRANKKIKERQLO 31

RESULT 20  
US-10-600-003-83  
Sequence 83, Application US/10600003  
Publication No. US20040197840A1  
GENERAL INFORMATION:  
APPLICANT: FOWLES, Dana M.  
BROACH, Jim  
MANFREDI, John  
KLEIN, Christine  
MURPHY, Andrew J.  
PAUL, Jeremy  
TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE AND COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/600,003  
FILING DATE: 18-Oct-2003  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/286,166  
FILING DATE: 05-APR-1999  
APPLICATION NUMBER: US 08/461,383  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/332,137  
FILING DATE: 13-OCT-1994  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CFI-012CP4B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-227-7400  
TELEFAX: 617-227-5941  
TELEFAX: 752806  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 65 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 83:  
US-10-600-003-83  
Query Match 65.6%; Score 42; DB 16; Length 65;  
Best Local Similarity 57.1%; Pred. No. 31;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ASREAKKQVERKALE 14  
Db 18 AQRANKKIKERQLO 31

RESULT 21  
US-10-424-599-223904  
Sequence 223904, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 223904  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(86)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_44214C.1.pap  
US-10-424-599-223904

Query Match 65.6%; Score 42; DB 15; Length 86;  
Best Local Similarity 57.1%; Pred. No. 41;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ASREAKKQVERKALE 14  
Db 6 AQRANKKIKERQLO 19

RESULT 22  
US-10-732-923-7618  
Sequence 7618, Application US/10732923  
Publication No. US20050108791A1  
GENERAL INFORMATION:  
APPLICANT: Edgerton, Michael D  
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
FILE REFERENCE: 38-15(52796)C  
CURRENT APPLICATION NUMBER: US/10/732,923  
CURRENT FILING DATE: 2003-12-10  
PRIOR APPLICATION NUMBER: 10/310,154  
PRIOR FILING DATE: 2002-12-04  
NUMBER OF SEQ ID NOS: 24149  
SEQ ID NO 7618  
LENGTH: 377  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-732-923-7618

Query Match 65.6%; Score 42; DB 17; Length 377;  
Best Local Similarity 57.1%; Pred. No. 1.9e+02;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ASREAKKQVERKALE 14  
Db 16 AQRANKKIKERQLO 29

RESULT 23  
US-09-952-680A-15

```
; Sequence 15, Application US/09952680A
; Publication No. US20030087239A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Mary
; APPLICANT: Epstein, David
; APPLICANT: Hamaguchi, No. US20030087239A1uko
; TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
; FILE REFERENCE: 23239-501
; CURRENT APPLICATION NUMBER: US/09/952,680A
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 60/223,454
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-952-680A-15
```

```
Query Match      65.6%; Score 42; DB 10; Length 379;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
OY      1 ASREAKKQVEKALE 14
      |||||:|:|:|:|
Db      18 AOREANKKIEKOLQ 31
```

```
RESULT 24
US-10-408-765A-105
; Sequence 105, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-105
```

```
Query Match      65.6%; Score 42; DB 16; Length 379;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
OY      1 ASREAKKQVEKALE 14
      |||||:|:|:|:|
Db      18 AOREANKKIEKOLQ 31
```

```
RESULT 25
US-10-215-982-15
; Sequence 15, Application US/10215982
; Publication No. US20040219523A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Martin
; APPLICANT: Epstein, David
; APPLICANT: Hamaguchi, Nobuko
; APPLICANT: Kurtz, Markus
; APPLICANT: Keefe, Tony
```

```
; APPLICANT: Wilson, Charles
; APPLICANT: Grate, Dilara
; APPLICANT: Marshall, Kristin
; APPLICANT: McCauley, Thomas
; APPLICANT: Kurtz, Jeffrey
; TITLE OF INVENTION: NUCLEIC ACID SENSOR MOLECULES AND METHODS OF USING SAME
; FILE REFERENCE: 23239-501 CIP
; CURRENT APPLICATION NUMBER: US/10/215,982
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/232,454
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 09/952,680
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 60/311,378
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/313,932
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/338,186
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/349,959
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/364,486
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/376,744
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/367,991
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/369,887
; PRIOR FILING DATE: 2002-04-04
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 372
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-215-982-15
```

```
Query Match      65.6%; Score 42; DB 16; Length 379;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
OY      1 ASREAKKQVEKALE 14
      |||||:|:|:|:|
Db      18 AOREANKKIEKOLQ 31
```

```
RESULT 26
US-10-732-923-7963
; Sequence 7963, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 7963
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-732-923-7963
```

```
Query Match      65.6%; Score 42; DB 17; Length 379;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;
```

```
OY      1 ASREAKKQVEKALE 14
      |||||:|:~|:~|:~|
Db      18 AOREANKKIEKOLQ 31
```

```
RESULT 27
US-10-732-923-8015
; Sequence 8015, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8015
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-8015

Query Match      65.6%; Score 42; DB 17; Length 379;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ASREKKQVERKALE 14
      |||||::|||:
Db      18 AGRANKKIKERKQLQ 31

RESULT 28
US-09-952-680A-16
; Sequence 16, Application US/09952680A
; Publication No. US20030087239A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Marty
; APPLICANT: Epstein, David
; APPLICANT: Hameguchi, No. US20030087239A1uko
; TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
; FILE REFERENCE: 23239-501
; CURRENT APPLICATION NUMBER: US/09/952,680A
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 60/232,454
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-952-680A-16

Query Match      65.6%; Score 42; DB 10; Length 380;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ASREKKQVERKALE 14
      |||||::|||:
Db      18 AGRANKKIKERKQLQ 31

RESULT 29
US-10-215-982-16
; Sequence 16, Application US/10215982
; Publication No. US20040219523A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Martin
; APPLICANT: Epstein, David
; APPLICANT: Hameguchi, Nobuko
; APPLICANT: Kurtz, Markus
; APPLICANT: Keefe, Tony
; APPLICANT: Wilson, Charles
; APPLICANT: Grate, Dilara
```

```
; APPLICANT: Marshall, Kristin
; APPLICANT: McCauley, Thomas
; APPLICANT: Kurtz, Jeffrey
; TITLE OF INVENTION: NUCLEIC ACID SENSOR MOLECULES AND METHODS OF USING SAME
; FILE REFERENCE: 23239-501 CIP
; CURRENT APPLICATION NUMBER: US/10/215,982
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/232,454
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 09/952,680
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 60/311,378
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/313,932
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/338,186
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/349,959
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/364,486
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/376,744
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/367,991
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/369,887
; PRIOR FILING DATE: 2002-04-04
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 372
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-215-982-16

Query Match      65.6%; Score 42; DB 16; Length 380;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ASREKKQVERKALE 14
      |||||::|||:
Db      18 AGRANKKIKERKQLQ 31

RESULT 30
US-10-732-923-8016
; Sequence 8016, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8016
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-8016

Query Match      65.6%; Score 42; DB 17; Length 380;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ASREKKQVERKALE 14
      |||||::|||:
Db      18 AGRANKKIKERKQLQ 31
```

```

RESULT 31
US-10-732-923-8055
: Sequence 8055, Application US/10732923
: Publication No. US20050108791A1
: GENERAL INFORMATION:
: APPLICANT: Edgerton, Michael D
: TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPIC
: FILE REFERENCE: 38-15(52796)C
: CURRENT APPLICATION NUMBER: US/10/732,923
: CURRENT FILING DATE: 2003-12-10
: PRIOR APPLICATION NUMBER: 10/310,154
: PRIOR FILING DATE: 2002-12-04
: NUMBER OF SEQ ID NOS: 24149
: SEQ ID NO 8055
: LENGTH: 380
: TYPE: PRT
: ORGANISM: Bob tauus
US-10-732-923-8055

```

Query Match	65.6%	Score 42;	DB 17;	Length 380;
Best Local Similarity	57.1%	Pred. No. 1.9e+02;		
Matches 8;	Conservative 3;	Mismatches 3;	Indels 0;	Gaps 0;

```
QY      1 ASREAKKQVEKALE 14
        | ||| | : : | | :
Db      18 AOREANKKIEKQLQ 31
```

```

RESULT 32
US-10-732-923-8022
: Sequence 8022, Application US/10/732923
: Publication No. US20050108791A1
: GENERAL INFORMATION:
: APPLICANT: Edgerton, Michael D
: TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
: FILE REFERENCE: 38-15(52796)C
: CURRENT APPLICATION NUMBER: US/10/732,923
: CURRENT FILING DATE: 2003-12-10
: PRIOR APPLICATION NUMBER: 10/310,154
: PRIOR FILING DATE: 2002-12-04
: NUMBER OF SEQ ID NOS: 24149
: SEQ ID NO 8022
: LENGTH: 384
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-732-923-8022

```

Query Match 65.6%; Score 42; DB 17; Length 384;  
 Best Local Similarity 57.1%;  
 Best Local Similarity Pred. No. 1.9e+02;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ASREAKKQVEKALE 14
    | ||| |::||| :
Db 7 AQREANKKIEKQLQ 20
```

```

1 RESULT 33
2 US-10-732-923-8054
3 ; Sequence 8054, Application US/107332923
4 ; Publication No. US20050108791A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Edgerton, Michael D
7 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
8 ; FILE REFERENCE: 38-15152796)C
9 ; CURRENT APPLICATION NUMBER: US/10/732,923
10 ; CURRENT FILING DATE: 2003-12-10
11 ; PRIOR APPLICATION NUMBER: 10/310,154
12 ; PRIOR FILING DATE: 2002-12-04
13 ; NUMBER OF SEQ ID NOS: 24149
14 ; SEQ ID NO 8054
15 ; LENGTH: 388
16 ; TYPE: PRT
17 ; ORGANISM: Bos taurus

```

US-10-732-923-8054

Query Match	65.6%	Score 42;	DB 17;	length 388;
Best Local Similarity	57.1%	Pred. No. 2e+02;		
Matches	8; Conservative	3; Mismatches	3; Indels	0; Gaps

Qy	1	ASREAKKQVEKALE	14
		:	:
Db	18	AQREANKKIEKQLQ	31

```

RESULT 34
US-09-952-680A-13
; Sequence 13. Application US/09952680A
; Publication No. US20030087239A1
GENERAL INFORMATION:
APPLICANT: Stanton, Marty
APPLICANT: Epstein, David
APPLICANT: Hamaguchi, No. US20030087239A1uko
TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
FILE REFERENCE: 23239-501
CURRENT APPLICATION NUMBER: US/09/952,680A
CURRENT FILING DATE: 2001-09-13
PRIOR APPLICATION NUMBER: 60/232,454
PRIOR FILING DATE: 2000-09-13
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-952-680A-13

```

Query Match	65.6%;	Score 42;	DB 10;	Length 394;
Best Local Similarity	57.1%;	Pred. No. 2e+02;		
Matches	8;	Conservative	3;	Mismatches 0;
				Gaps 0;

```
QY      1 ASREAKKQVEKALE 14
        | ||| |::|| |
Db      18 AQRANKKIEKQLQ 31
```

```

RESULT 35
US-09-963-131-192
; Sequence 192, Application US/09963131
; Publication No. US2003022460A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn Skou
; APPLICANT: Sorensen, Annette Balle
; APPLICANT: Hernandez, Javier Martin
; APPLICANT: Nielsen, Anne Ahlmann
; APPLICANT: Moving, Helle
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR LYMPHOMA AND LEUKEMIA
; FILE REFERENCE: 52945200323
; CURRENT APPLICATION NUMBER: US/09/963,131
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/905,390
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/905,491
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 192
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-963-131-192

```

Query Match	65.6%	Score 42;	DB 10;	Length 394;
Best Local Similarity	57.1%	Pred. No. 2e+02;		
Matches	8;	Conservative	3;	Mismatches 3;
				Indels 0;
				Gaps 0



Qy 1 ASREAKKQVEKALE 14  
| | | | : : : | :  
Db 18 AOREANKKIEKQLO 31

RESULT 36  
US-09-963-131-194  
; Sequence 194, Application US/09963131  
; Publication No. US20030224460A1  
; GENERAL INFORMATION:  
; APPLICANT: Pedersen, Finn Skou  
; APPLICANT: Sorensen, Annette Baile  
; APPLICANT: Hernandez, Javier Martin  
; APPLICANT: Nielsen, Anne Ahlmann  
; APPLICANT: Moring, Helge  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR LYMPHOMA AND LEUKEMIA  
; FILE REFERENCE: 529452000323  
; CURRENT APPLICATION NUMBER: US/09/963.131  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: US 09/668,644  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: US 09/905,390  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 09/905,491  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 215  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 194  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-963-131-194

Query Match 65.6%; Score 42; DB 10; Length 394;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
| | | | : : : | :  
Db 18 AOREANKKIEKQLO 31

RESULT 37  
US-10-116-275-187  
; Sequence 187, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Eilan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Breiden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 187  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-187

Query Match 65.6%; Score 42; DB 15; Length 394;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
| | | | : : : | :  
Db 18 AOREANKKIEKQLO 31

Db 18 AOREANKKIEKQLO 31

RESULT 38  
US-10-352-843-18  
; Sequence 18, Application US/10352843  
; Publication No. US20040014135A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Lisa  
; APPLICANT: Kindt, Rachel  
; APPLICANT: Kopczynski, Jenny  
; APPLICANT: Doherty, Stephen  
; APPLICANT: Cockell, Mark  
; APPLICANT: Ramanathan, Chandra  
; APPLICANT: Lodge, Nicholas  
; APPLICANT: Fitzgerald, Kevin  
; APPLICANT: Stouch, Terry  
; TITLE OF INVENTION: MOLECULES THAT MODULATE G(ALPHA)q ACTIVITY AND METHODS OF  
; FILE REFERENCE: 5624-277-999  
; CURRENT APPLICATION NUMBER: US/10/352,843  
; CURRENT FILING DATE: 2003-01-27  
; PRIOR APPLICATION NUMBER: US 60/352720  
; PRIOR FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: G-protein of the invention  
US-10-352-843-18

Query Match 65.6%; Score 42; DB 15; Length 394;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
| | | | : : : | :  
Db 18 AOREANKKIEKQLO 31

RESULT 39  
US-10-215-982-13  
; Sequence 13, Application US/10215982  
; Publication No. US20040219523A1  
; GENERAL INFORMATION:  
; APPLICANT: Stanton, Martin  
; APPLICANT: Epstein, David  
; APPLICANT: Hamaguchi, Nobuko  
; APPLICANT: Kurz, Markus  
; APPLICANT: Keefe, Tony  
; APPLICANT: Wilson, Charles  
; APPLICANT: Grate, Dilara  
; APPLICANT: Marshall, Kristin  
; APPLICANT: McCauley, Thomas  
; TITLE OF INVENTION: NOCISIC ACID SENSOR MOLECULES AND METHODS OF USING SAME  
; FILE REFERENCE: 23239-501 CIP  
; CURRENT APPLICATION NUMBER: US/10/215,982  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 60/232,454  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: 09/952,680  
; PRIOR FILING DATE: 2001-09-13  
; PRIOR APPLICATION NUMBER: 60/311,378  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/313,932  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/338,186  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/349,959



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: June 13, 2005, 20:28:35 / Search time 15.6279 Seconds  
(without alignments)  
86.194 Million cell updates/sec

Title: US-10-706-275-1  
Perfect score: 64  
Sequence: 1 ASREAKKQVERALE 14

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	388	2 A49545	plasma10gen-bindin
2	64	100.0	408	2 S30283	protein M precursor
3	64	100.0	436	2 S30284	M protein precursor
4	64	100.0	454	2 S43556	plasma10gen-bindin
5	64	100.0	472	2 S43554	plasma10gen-bindin
6	64	100.0	483	2 A26297	M6 protein - Strept
7	64	100.0	484	2 S35401	M1 protein precursor
8	64	100.0	484	2 S46489	M1 protein precursor
9	64	100.0	484	2 S34978	M1.1 protein precursor
10	64	100.0	492	2 A28616	M5 protein precursor
11	64	100.0	501	2 A44643	M protein precursor
12	64	100.0	528	2 S57835	Igg-binding protei
13	64	100.0	532	2 S54871	M protein - Strept
14	64	100.0	539	2 A28549	M24 protein precursor
15	64	100.0	564	2 A60115	M protein precursor
16	64	100.0	587	2 UCI419	Fc gamma (Igg) rec
17	45	70.3	546	2 T33382	hypothetical prote
18	44	68.8	149	2 T48000	conserved hypothet
19	44	68.8	423	2 T48000	hypothetical prote
20	43	67.2	104	1 H64327	conserved hypothet
21	43	67.2	168	2 PH0139	M protein-like mol
22	43	67.2	365	2 B54128	Fc-binding protein
23	43	67.2	384	2 S49550	M-like protein emm
24	43	67.2	386	2 S05568	Iga receptor precu
25	43	67.2	386	2 S54858	M protein precursor
26	43	67.2	389	2 A43715	M49 protein precursor
27	43	67.2	402	2 S37046	Iga receptor - Str
28	43	67.2	407	2 S23325	M2 protein precursor
29	42	65.6	91	2 A46685	GTP-binding regula

30	42	65.6	377	1 RGM5A1	GRF-binding regula
31	42	65.6	380	1 RGHU1	GRF-binding regula
32	42	65.6	394	1 RGHVAE	GRF-binding regula
33	42	65.6	394	1 RGM5A2	GRF-binding regula
34	42	65.6	394	1 RGHVA2	GRF-binding regula
35	42	65.6	394	1 RGHVA2	GRF-binding regula
36	42	65.6	394	1 RGHVA2	GRF-binding regula
37	42	65.6	395	1 RGHUA2	GRF-binding regula
38	42	65.6	397	1 RGP5A2	GRF-binding regula
39	42	65.6	493	2 T22180	hypothetical prote
40	41	64.1	132	2 T06930	H+-transporting tw
41	41	64.1	269	2 A86641	hypothetical prote
42	41	64.1	1328	2 T23007	hypothetical prote
43	40	62.5	127	2 S78352	ribosomal protein
44	40	62.5	157	2 H84093	small protein B BH
45	40	62.5	377	2 S52537	emm L 15 protein -

## ALIGNMENTS

## RESULT 1

A49545  
N:Alternate names: plasma10gen-binding M-like protein (Pd 53)  
C:Species: Streptococcus pyogenes  
C>Date: 23-Mar-1995 #sequence revision 05-Apr-1995 #text\_change 09-Jul-2004  
C:Accession: A49545; S61084; S60829; S70459; S32619  
R:Berge, A.; Sjoeborg, U.  
J. Biol. Chem. 268, 25417-25424, 1993  
A>Title: PAM, a novel plasma10gen-binding protein from Streptococcus pyogenes.  
A:Reference number: A49545; MUID:94064605; PMID:8244975  
A:Accession: A49545  
A:Molecule type: DNA  
A:Residues: 1-388 <BER>  
A:Cross-references: UNIPROT:P49054; EMBL:Z22219; NID:G288978; PIDN:CAA80222.1; PID:99408  
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
submitted to the EMBL Data Library, July 1994  
A:Description: Noncongruent relationships between variation in emm1 gene sequences and t  
A:Reference number: S61072  
A:Accession: S61084  
A:Molecule type: DNA  
A:Residues: 13-96 <MHA>  
A:Cross-references: EMBL:U11975; NID:G533627; PIDN:AAA9591.1; PID:G1235829  
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A>Title: Non-congruent relationships between variation in emm gene sequences and the pop  
A:Reference number: S60784; MUID:95198537; PMID:7891551  
A:Accession: S60829  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 17-77 <MW>  
A:Cross-references: EMBL:U11975  
R:Carlsson Wistedt, A.; Ringdahl, U.; Mueller-Eberhard, W.; Sjoeborg, U.  
Mol. Microbiol. 18, 569-578, 1995  
A>Title: Identification of a plasma10gen-binding motif in PAM, a bacterial surface prote  
A:Reference number: S70457; MUID:96342385; PMID:8746039  
A:Accession: S70459  
A:Molecule type: DNA  
A:Residues: 30-162 <CAR>  
C:Superfamily: M5 protein  
C:Keywords: cell wall  
F11-29/Domains: signal sequence (fragment) #status predicted <SIG>  
F30-388/Product: plasma10gen-binding protein PAM #status predicted <MAT>

Query Match 100.0%; Score 64; DB 2; Length 388;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKQVERALE 14  
|||||

Db 266 ASREAKKOVERKALE 279

RESULT 2

S30283  
protein M precursor - Streptococcus pyogenes (serotype M41)  
C/Species: Streptococcus pyogenes  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C/Accession: S30283; S29680  
R;Podbielski, A.  
Mol. Gen. Genet. 237, 287-300, 1993  
A;Title: Three different types of organization of the vir regulon in group A streptococci  
A;Reference number: S30283; MUID:93204905; PMID:8455563  
A;Accession: S30283  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-408 <POD1>  
A;Cross-references: UNIPROT:Q54837; EMBL:X58178  
R;Podbielski, A.; Melzer, B.  
submitted to the EMBL Data Library, February 1991  
A;Description: Pcr mediated cloning and sequencing of group A streptococcal emm41/52 (ex  
A;Reference number: S29680  
A;Accession: S29680  
A;Molecule type: DNA  
A;Residues: 1-230,'N',232-371,'R',373-408 <POD2>  
A;Cross-references: EMBL:X58178; NID:947362; PIDN:CAA41167.1; PID:947363  
C;Genetics:  
A;Gene: emm  
C;Superfamily: M5 protein  
C;Keywords: transmembrane protein  
F;1-41/Domain: signal sequence #status predicted <SIG>  
F;42-408/Product: M protein #status predicted <MAT>  
F;383-401/Domain: transmembrane #status predicted <TMM>

Query Match 100.0%; Score 64; DB 2; Length 408;  
Best Local Similarity 100.0%; Pred. No. 0.02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 259 ASREAKKOVERKALE 272

RESULT 3

S30284  
M protein precursor - Streptococcus pyogenes (serotype M52)  
C/Species: Streptococcus pyogenes  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C/Accession: S30284; S29681  
R;Podbielski, A.  
Mol. Gen. Genet. 237, 287-300, 1993  
A;Title: Three different types of organization of the vir regulon in group A streptococci  
A;Reference number: S30283; MUID:93204905; PMID:8455563  
A;Accession: S30284  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-436 <POD1>  
A;Cross-references: UNIPROT:Q54839; EMBL:X58179  
R;Podbielski, A.; Melzer, B.  
submitted to the EMBL Data Library, February 1991  
A;Description: Pcr mediated cloning and sequencing of group A streptococcal emm41/52 (ex  
A;Reference number: S29680  
A;Accession: S29680  
A;Molecule type: DNA  
A;Residues: 1-216,'N',218-436 <POD2>  
A;Cross-references: EMBL:X58179; NID:947364; PIDN:CAA41168.1; PID:947365  
C;Genetics:  
A;Gene: emm  
C;Superfamily: M5 protein  
C;Keywords: transmembrane protein  
F;1-11/Domain: signal sequence #status predicted <SIG>  
F;42-436/Product: M protein #status predicted <MAT>  
F;411-429/Domain: transmembrane #status predicted <TMM>

Query Match 100.0%; Score 64; DB 2; Length 436;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 287 ASREAKKOVERKALE 300

RESULT 4

S4356  
plasmidogen-binding protein MLC36 - Streptococcus sp. (fragment)  
C/Species: Streptococcus sp.  
C/Date: 14-Sep-1994 #sequence\_revision 03-Nov-1995 #text\_change 26-Aug-1999  
C/Accession: S45598; S4356  
R;Ben Nasr, A.; Wistedt, A.; Ringdahl, U.; Sjoebirg, U.  
Eur. J. Biochem. 222, 267-276, 1994  
A;Title: Streptokinase activates plasmidogen bound to human group C and G streptococci  
A;Reference number: S45598; MUID:94291620; PMID:8020466  
A;Accession: S45598  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-454 <BE2>  
A;Cross-references: EMBL:Z32677; NID:9474767; PIDN:CAA83588.1; PID:9474768  
C;Superfamily: M5 protein

Query Match 100.0%; Score 64; DB 2; Length 454;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 332 ASREAKKOVERKALE 345

RESULT 5

S4354  
plasmidogen-binding protein MLC72 - Streptococcus sp. (fragment)  
C/Species: Streptococcus sp.  
A;Variety: group G  
C/Date: 07-Sep-1994 #sequence\_revision 03-Nov-1995 #text\_change 20-Jun-2000  
C/Accession: S45599; S4354  
R;Ben Nasr, A.; Wistedt, A.; Ringdahl, U.; Sjoebirg, U.  
Eur. J. Biochem. 222, 267-276, 1994  
A;Title: Streptokinase activates plasmidogen bound to human group C and G streptococci  
A;Reference number: S45598; MUID:94291620; PMID:8020466  
A;Accession: S45599  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-472 <BE2>  
A;Cross-references: EMBL:Z32678; NID:9474769; PIDN:CAA83589.1; PID:94333838  
C;Superfamily: M5 protein

Query Match 100.0%; Score 64; DB 2; Length 472;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 350 ASREAKKOVERKALE 363

RESULT 6

A26297  
M6 protein - Streptococcus pyogenes  
C/Species: Streptococcus pyogenes  
C/Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 09-Jul-2004  
C/Accession: A26297  
R;Hollingshead, S.K.; Flechett, V.A.; Scott, J.R.  
J. Biol. Chem. 263, 1677-1686, 1986  
A;Title: Complete nucleotide sequence of type 6 M protein of the group A streptococcus.  
A;Reference number: A26297; MUID:86111835; PMID:3511046

A/Accession: A26297  
A/Molecule type: DNA  
A/Residues: 1-483 <HOU>  
A/Cross-references: UNIPROT:P08089; GB:M1J38; GB:M1J45; NID:G153699; PIDN:AAA26920.1;  
C/Genetics:  
A/Gene: emm6  
C/Superfamily: M5 protein  
C/Keywords: coiled coil; transmembrane protein

Query Match 100.0%; Score 64; DB 2; Length 483;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASREAKKQVEKALE 14  
|||||  
Db 334 ASREAKKQVEKALE 347

## RESULT 7

S35401  
M1 protein precursor - Streptococcus pyogenes

C/Species: Streptococcus pyogenes

A/Variety: serotype M1

C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C/Accession: S35401; S61074; S60784

R/PdbId:1sk1, A.  
submitted to the EMBL Data Library, September 1991

A/Reference number: S35401

A/Accession: S35401

A/Molecule type: DNA

A/Residues: 1-484 <PDB>

A/Cross-references: UNIPROT:Q10372; EMBL:X62131; NID:G311757; PIDN:CAA44062.1; PID:G3117

R/Hatmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.

A/Title: Non-congruent relationships between variation in emm1 gene sequences and t

A/Description: Noncongruent relationships between variation in emm1 gene sequences and t

A/Reference number: S61072

A/Accession: S61074

A/Molecule type: DNA

A/Status: preliminary

A/Residues: 16-94 <MHA>

A/Cross-references: EMBL:U11940; NID:G533557; PIDN:AAA9556.1; PID:G533558

R/Hatmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A/Title: Non-congruent relationships between variation in emm gene sequences and the po

A/Reference number: S60784; MUID:95198537; PMID:7891551

A/Accession: S60784

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 29-89 <WH2>

A/Cross-references: EMBL:U11940

C/Genetics:

A/Gene: emm1

C/Superfamily: M5 protein

Query Match 100.0%; Score 64; DB 2; Length 484;

Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASREAKKQVEKALE 14  
|||||

Db 335 ASREAKKQVEKALE 348

## RESULT 8

S46489  
M1 protein precursor - Streptococcus pyogenes

C/Species: Streptococcus pyogenes

C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C/Accession: S46489; S46490

R/Akesson, P.; Schmidt, K.H.; Cooney, J.; Bjorck, L.

Biochem. J. 300, 877-886, 1994

A/Title: M1 protein and protein H: IgGFC- and albumin-binding streptococcal surface prot

A/Accession: S46489  
A/Molecule type: DNA  
A/Residues: 1-484 <AXE>  
A/Cross-references: UNIPROT:O05464; UNIPROT:Q10372; UNIPROT:Q99XV0

A/Experimental source: strain 40/58, serotype M1

A/Accession: S46490

A/Molecule type: protein

A/Residues: 42-51 <AKW>

A/Experimental source: strain 40/58, serotype M1

C/Genetics:

A/Gene: emm1

C/Superfamily: M5 protein

C/Keywords: transmembrane protein

F:1-41/Domain: signal sequence #status predicted <SIG>

F:42-484/Product: M1 protein #status experimental <MAT>

F:459-477/Domain: transmembrane #status predicted <TM>

Query Match 100.0%; Score 64; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASREAKKQVEKALE 14  
|||||

Db 335 ASREAKKQVEKALE 348

## RESULT 9

S34978  
M1.1 protein precursor - Streptococcus pyogenes

C/Species: Streptococcus pyogenes

C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C/Accession: S34978; S31966

R/Haugh, M.P.; Podbielski, A.; Huegl, S.; Cleary, P.P.

Mol. Microbiol. 9, 981-991, 1993

A/Title: Nucleotide substitutions and small-scale insertion produce size and antigenic v

A/Reference number: S34978; MUID:93360826; PMID:8355619

A/Accession: S34978

A/Molecule type: DNA

A/Residues: 1-484 <HAR>

A/Cross-references: UNIPROT:Q05464; EMBL:Z21845; NID:G49401; PIDN:CAA79893.1; PID:G49402

C/Genetics:

A/Gene: emm1.1

C/Superfamily: M5 protein

C/Keywords: membrane protein

F:1-42/Domain: signal sequence #status predicted <SIG>

F:43-484/Product: M1.1 protein #status predicted <MAT>

Query Match 100.0%; Score 64; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASREAKKQVEKALE 14  
|||||

Db 335 ASREAKKQVEKALE 348

## RESULT 10

A28616  
M5 protein precursor - Streptococcus pyogenes

C/Species: Streptococcus pyogenes

A/Variety: serotype M5

C/Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 09-Jul-2004

C/Accession: A28616; S60787

R/Miller, L.; Gray, L.; Beachey, E.; Kehoe, M.

J. Biol. Chem. 263, 5668-5673, 1988

A/Title: Antigenic variation among group A streptococcal M proteins. Nucleotide sequence

A/Reference number: A28616; MUID:88186881; PMID:3281944

A/Accession: A28616

A/Molecule type: DNA

A/Residues: 1-492 <MIL>

A/Cross-references: UNIPROT:P02977; GB:M20374; NID:G153812; PIDN:AAA26976.1; PID:G153813

R/Hatmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A>Title: Non-congruent relationships between variation in emm gene sequences and the po  
 A:Reference number: S60784; MUID:95198537; PMID:7891551  
 A:Accession: S60787  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: DNA  
 A:Residues: 30-89 <MHA>  
 C:Genetics:  
 A:Gene: smps  
 C:Superfamily: M5 protein  
 C:Keywords: coiled coil; transmembrane protein  
 F:1-42/Domain: signal sequence #status predicted <SIG>  
 F:43-492/Product: M5 protein #status predicted <MAT>

Query Match 100.0%; Score 64; DB 2; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 343 ASREAKKQVEKALE 356

RESULT 11  
 A44643  
 M protein precursor - Streptococcus pyogenes (serotype M57) (fragment)  
 C:Species: Streptococcus pyogenes  
 A:Variety: serotype M57  
 C:Date: 26-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
 C:Accession: A44643; S60833  
 R:Manjula, B.N.; Khandke, K.M.; Fairwell, T.; Relif, W.A.; Sripkrash, K.S.  
 J. Protein Chem. 10, 369-384, 1991  
 A>Title: Heptad motifs within the distal subdomain of the coiled-coil rod region of M pr  
 her: nucleotide sequence of the M57 gene and relation of the dedu.  
 A:Reference number: A44643; MUID:9243933; PMID:1781883  
 A:Accession: A44643  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-501 <MAN>  
 A:Experimental source: type M57, strain A995  
 A>Note: sequence inconsistent with nucleotide translation  
 A>Note: sequence extracted from NCBI backbone (NCBIN:83737, NCBIP:83738)  
 A>Note: parts of this sequence were confirmed by peptide sequencing  
 R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kenoe, M.A.  
 Mol. Microbiol. 14, 619-631, 1994  
 A>Title: Non-congruent relationships between variation in emm gene sequences and the po  
 A:Reference number: S60784; MUID:95198537; PMID:7891551  
 A:Accession: S60833  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 14-95 <MHA>  
 A:Cross-references: EMBL:U11971  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
 C:Superfamily: M5 protein  
 C:Keywords: coiled coil; dimer

Query Match 100.0%; Score 64; DB 2; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 0.025;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 364 ASREAKKQVEKALE 377

RESULT 12  
 S57835  
 IGG-binding protein emml precursor - Streptococcus pyogenes (strain 64/14)  
 N:Alternate names: IGG-binding protein type IIA; type IIA immunoglobulin G-binding prote  
 C:Species: Streptococcus pyogenes  
 A:Variety: strain 64/14  
 C:Date: 28-Nov-1995 #sequence\_revision 09-Mar-1996 #text\_change 09-Jul-2004  
 C:Accession: S57835; S58931  
 R:Boyle, M.D.P.; Hawliltzy, J.; Raeder, R.; Podbielski, A.

Infect. Immun. 62, 1336-1347, 1994  
 A>Title: Analysis of genes encoding two unique type IIA immunoglobulin G-binding proteins  
 A:Reference number: S57834; MUID:94178942; PMID:8132341  
 A:Accession: S57835  
 A:Molecule type: DNA  
 A:Residues: 1-528 <BOY>  
 A:Cross-references: UNIPROT:Q54843; EMBL:X72932  
 A:Experimental source: strain 64/14  
 A>Note: the authors translated the codons CTTAAA for residue 52 and 53 as Arg  
 R:Podbielski, A.  
 submitted to the EMBL Data Library, March 1993  
 A:Reference number: S58931  
 A:Accession: S58931  
 A:Molecule type: DNA  
 A:Residues: 1-46, 'E', 48-52, 'E', 54-528 <POD>  
 A:Cross-references: EMBL:X72932; NID:9507128; PIDN:CAA51437.1; PID:9507130  
 A:Experimental source: strain 64/14  
 C:Genetics:  
 A:Gene: emml  
 C:Superfamily: M5 protein  
 F:1-41/Domain: signal sequence #status predicted <SIG>  
 F:42-528/Product: type IIA immunoglobulin G-binding protein emml #status predicted <MAT>

Query Match 100.0%; Score 64; DB 2; Length 528;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 379 ASREAKKQVEKALE 392

RESULT 13  
 S54871  
 M protein - Streptococcus sp.  
 C:Species: Streptococcus sp.  
 C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
 C:Accession: S54871  
 R:Podbielski, A.; Melzer, B.  
 submitted to the EMBL Data Library, June 1991  
 A:Reference number: S54871  
 A:Accession: S54871  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-532 <POD>  
 A:Cross-references: UNIPROT:Q55098; EMBL:X60097; NID:9840905; PIDN:CAA42693.1; PID:984091  
 C:Superfamily: M5 protein

Query Match 100.0%; Score 64; DB 2; Length 532;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 383 ASREAKKQVEKALE 396

RESULT 14  
 A28549  
 M24 protein precursor - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes  
 A:Variety: serotype M24  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
 C:Accession: A28549; S60802  
 R:Mouw, A.R.; Beachey, E.H.; Burdett, V.  
 U. Bacteriol. 170, 676-684, 1988  
 A>Title: Molecular evolution of streptococcal M protein: cloning and nucleotide sequence  
 A:Reference number: A28549; MUID:88115166; PMID:3276665  
 A:Accession: A28549  
 A:Molecule type: DNA  
 A:Residues: 1-539 <MOU>  
 A:Cross-references: UNIPROT:P12379; GB:MI9031; NID:9153616; PIDN:AAA26874.1; PID:9153617  
 R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kenoe, M.A.

Mol. Microbiol. 14, 619-631, 1994  
A/Title: Non-congruent relationships between variation in emm gene sequences and the por  
A/Reference number: S60784; MUID:95198537; PMID:7891551  
A/Accession: S60802  
A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A/Molecule type: DNA  
A/Residues: 30-89 <WHA>  
A/Superfamily: M5 protein  
C/Keywords: coiled coil; transmembrane protein

Query Match 100.0%; Score 64; DB 2; Length 539;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 390 ASREAKKOVERKALE 403

RESULT 15  
M protein precursor - Streptococcus pyogenes (serotype M12) (fragment)  
C/Species: Streptococcus pyogenes  
A/Variety: serotype M12  
C/Date: 08-Dec-1992 #sequence\_revision 08-Dec-1992 #text\_change 09-Jul-2004  
A/Accession: A40174; A60115; S39887; S61072; S60793  
R/Robbins, J.C.; Spangler, J.G.; Jones, S.J.; Simpson, W.J.; Cleary, P.P.  
J. Bacteriol. 169, 5633-5640, 1987  
A/Title: Streptococcus pyogenes type 12 M protein gene regulation by upstream sequences.  
A/Reference number: A40174; MUID:88058777; PMID:2445730  
A/Accession: A40174  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-564 <ROB>  
A/Cross-references: UNIPROT:P19401; GB:M18269; NID:9153543; PIDN:AAA88573.1; PID:9153544  
R/Kraus, W.; Seyer, J.M.; Beachey, E.H.  
Infect. Immun. 57, 2457-2461, 1989  
A/Title: Vimentin-cross-reactive epitope of type 12 streptococcal M protein.  
A/Reference number: A60115; MUID:89307564; PMID:2473037  
A/Accession: A60115  
A/Molecule type: protein  
A/Residues: 42-54 <KRA>  
R/Chen, C.; Bormann, N.; Cleary, P.P.  
Mol. Gen. Genet. 241, 685-693, 1993  
A/Title: Vrrr and Mrr are homologous trans-acting regulators of M protein and Csa peptid  
A/Reference number: S39886; MUID:94088463; PMID:7505389  
A/Accession: S39887  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-15 <CHB>  
R/Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
submitted to the EMBL Data Library, July 1994  
A/Description: Noncongruent relationships between variation in emm1 gene sequences and  
A/Reference number: S61072  
A/Accession: S61072  
A/Molecule type: DNA  
A/Residues: 13-111 <WHA>  
A/Cross-references: EMBL:U11937; NID:9533551; PIDN:AAA9553.1; PID:91235807  
R/Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A/Title: Non-congruent relationships between variation in emm gene sequences and the por  
A/Reference number: S60784; MUID:95198537; PMID:7891551  
A/Accession: S60793  
A/Status: nucleic acid sequence not shown  
A/Molecule type: DNA  
A/Residues: 29-89 <WHA>  
A/Cross-references: EMBL:U11937  
C/Genetics:  
A/Gene: emm12  
C/Superfamily: M5 protein  
C/Keywords: transmembrane protein  
F/1-41/Domain: signal sequence #status predicted <SIG>  
F/42-564/Product: M protein (fragment) #status predicted <MAT>

Query Match 100.0%; Score 64; DB 2; Length 564;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 427 ASREAKKOVERKALE 440

RESULT 16  
Fc gamma (19G) receptor II precursor - Streptococcus sp.  
JCI1419  
N/Alternate names: fcγv protein  
C/Species: Streptococcus sp.  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
A/Accession: JCI1419; S17354  
R/Smirnov, O.Y.; Deneslyuk, A.I.; Zakharov, M.V.; Abramov, V.M.; Zav'yalov, V.P.  
Gene 120, 27-32, 1992  
A/Title: Protein V, a novel type-II IgG receptor from Streptococcus sp.: Sequence, homo  
A/Reference number: JCI1419; MUID:93013016; PMID:1398120  
A/Accession: JCI1419  
A/Molecule type: DNA  
A/Residues: 1-587 <SMI>  
A/Cross-references: UNIPROT:Q53112; EMBL:X62467; NID:947562; PIDN:CAA44324.1; PID:947563  
A/Experimental source: strain 22/58 'Valente'  
C/Genetics:  
A/Gene: fcγv  
C/Superfamily: M5 protein  
C/Keywords: duplication; immunoglobulin receptor  
F/1-41/Domain: signal sequence #status predicted <SIG>  
F/42-587/Product: IgG Fc receptor II #status predicted <MAT>  
F/234-268/Region: 35-residue repeat A  
F/269-303/Region: 35-residue repeat A  
F/304-338/Region: 35-residue repeat A  
F/339-373/Region: 35-residue repeat A  
F/374-408/Region: 35-residue repeat B  
F/416-450/Region: 35-residue repeat B

Query Match 100.0%; Score 64; DB 2; Length 587;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 438 ASREAKKOVERKALE 451

RESULT 17  
hypochemical protein F53G2.7 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
A/Accession: T32382  
R/Becker, M.; Wohldmann, P.; Biewald, T.  
submitted to the EMBL Data Library, September 1997  
A/Description: The sequence of C. elegans cosmid F53G2.  
A/Reference number: Z21158  
A/Accession: T32382  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-546 <BEC>  
A/Cross-references: UNIPROT:O1245; EMBL:AF025464; PIDN:AAW71017.1; GSPDB:GN00020; CESP:  
A/Experimental source: strain Bristol N2; clone F53G2  
C/Genetics:  
A/Gene: CESP:F53G2.7  
A/Map position: 2  
A/Introns: 125/2; 172/2; 241/2; 236/2; 344/2; 374/3; 488/3; 519/3

Query Match 70.3%; Score 45; DB 2; Length 546;  
Best Local Similarity 75.0%; Pred. No. 22;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 REAKKOVERKALE 14  
|||:|:|:|  
Db 406 REKKOIEKALE 417

RESULT 18  
C81248  
conserved hypothetical protein NMB0013 [imported] - Neisseria meningitidis (strain MCS8)  
C:Species: Neisseria meningitidis  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: C81248  
R:Retcelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.B.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.  
A:Reference number: A81000; MUID:20175755; PMID:10710307  
A:Accession: C81248  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-149 <TE>  
A:Cross-references: UNIPROT:Q9K1Q7; GB:A8002359; GB:A8002098; NID:97225225; PIDN:AAF4045  
A:Experimental source: serogroup B, strain MCS8  
C:Genetics:  
A:Gene: NMB0013  
C:Superfamily: Campylobacter jejuni probable integral membrane protein Cj0830

Query Match 68.8%; Score 44; DB 2; Length 149;  
Best Local Similarity 66.7%; Pred. No. 9.4;  
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 SREAKKOVERKAL 13  
|||:|:|:|  
Db 38 SREARREVERKAM 49

RESULT 19  
T48000  
hypothetical protein T17J13.40 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T48000  
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24482  
A:Accession: T48000  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-423 <RIE>  
A:Cross-references: UNIPROT:Q9M1R6; EMBL:AL138651  
A:Experimental source: cultivar Columbia; BAC clone T17J13  
C:Genetics:  
A:Map position: 3  
A:Introns: 64/3; 84/3; 140/3; 201/3; 245/2; 301/3; 345/1  
A>Note: T17J13.40

Query Match 68.8%; Score 44; DB 2; Length 423;  
Best Local Similarity 81.8%; Pred. No. 25;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 EAKKOVERKALE 14  
|||:|:|:|  
Db 335 ESQKOVERKALE 345

RESULT 20  
H64327  
conserved hypothetical protein MJ0223 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: H64327  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodak, A.;  
son, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurest, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:868087  
A:Accession: H64327  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-104 <BUL>  
A:Cross-references: UNIPROT:Q57676; GB:U67478; GB:L77117; NID:g1590958; PIDN:AA898215.1;  
C:Genetics:  
A:Map position: REV214474-214160  
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0223

Query Match 67.2%; Score 43; DB 1; Length 104;  
Best Local Similarity 64.3%; Pred. No. 9.5;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ASREAKKOVERKALE 14  
|||:|:|:|  
Db 63 ABEAKKAEKALE 76

RESULT 21  
PH0139  
M protein-like molecule class II - Streptococcus sp. (group A) (fragment)  
C:Species: Streptococcus sp.  
C>Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 26-Aug-1999  
C:Accession: PH0139; S23067  
R:Beessen, D.E.; Fischetti, V.A.  
J. Exp. Med. 172, 1757-1764, 1990  
A>Title: Differentiation between two biologically distinct classes of group A streptococ-  
Rh4)  
A:Reference number: PH0139; MUID:91079780; PMID:2258705  
A:Accession: PH0139  
A:Molecule type: DNA  
A:Residues: 1-168 <BBS>  
A:Cross-references: GB:X6398; NID:947367; PIDN:CAA39808.1; PID:947368  
A:Experimental source: strain T2/44/Rd4  
C:Genetics:  
A:Gene: emm12.1  
C:Superfamily: M5 protein

Query Match 67.2%; Score 43; DB 2; Length 168;  
Best Local Similarity 76.9%; Pred. No. 15;  
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ASREAKKOVERKAL 13  
|||:|:|:|  
Db 141 ASREAKKVEADL 153

RESULT 22  
BS4128  
Fc-binding protein Sir22 precursor - Streptococcus pyogenes  
N:Alternate names: M protein  
C:Species: Streptococcus pyogenes  
A:Variety: serotype M22  
C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: BS4128; S60800  
R:Stenberg, L.; O'Toole, P.W.; Westecky, J.; Lindahl, G.  
J. Biol. Chem. 269, 13458-13464, 1994  
A>Title: Molecular characterization of protein Sir, a streptococcal cell surface protein  
A:Reference number: A54128; MUID:94230454; PMID:8175778  
A:Accession: BS4128  
A:Molecule type: DNA  
A:Residues: 1-365 <STB>  
A:Cross-references: UNIPROT:Q54901; GB:X75750; NID:9473162; PIDN:CAA53379.1; PID:9473164  
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A>Title: Non-congruent relationships between variation in emm gene sequences and the pop  
A:Reference number: S60784; MUID:95198537; PMID:7891551



A/Accession: S60800  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 29-109 <HNA>  
 A/Cross-references: EMBL:U11955; NID:G533587; PIDN:AAA9571.1; PID:G1235819  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
 C/Comment: The amino-terminal half of this protein mediates binding to both IgG and IgA  
 C/Genetics:  
 A/Gene: sir22  
 C/Superfamily: M5 protein  
 C/Keywords: cell wall  
 F:1-41/Domain: signal sequence #status predicted <SIG>  
 F:42-365/Product: Fe-binding protein Sir22 #status experimental <MAT>

Query Match 67.2%; Score 43; DB 2; Length 365;  
 Best Local Similarity 76.9%; Pred. No. 32;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASREAKKQVEAL 13  
 DB 236 ASREAKKKVEADL 248

RESULT 23  
 S49550  
 M-Like protein emm19 - Streptococcus pyogenes (strain 71-683)  
 C/Species: Streptococcus pyogenes  
 A/Variety: strain 71-683  
 C/Date: 09-Jun-1994 #sequence\_revision 06-Sep-1996 #text\_change 09-Jul-2004  
 C/Accession: S49550; S60141  
 R/Podbielski, A.; Hawlitzky, J.; Pack, T.D.; Flisbort, A.; Boyle, M.D.P.  
 Mol. Microbiol. 12, 725-736, 1994  
 A/Title: A group A streptococcal Ecm protein potentially resulting from intergenomic recombination  
 A/Reference number: S49551; MUID:94328925; PMID:8052125  
 A/Accession: S49550  
 A/Molecule type: DNA  
 A/Residues: 1-384 <POD>  
 A/Cross-references: UNIPROT:Q54841; EMBL:X72752; NID:G507126; PIDN:CAA51281.1; PID:G5071  
 A/Note: the authors translated the codon CAA for residue 247 as Lys, GAC for residue 266  
 C/Genetics:  
 A/Gene: emm19  
 C/Superfamily: M5 protein

Query Match 67.2%; Score 43; DB 2; Length 384;  
 Best Local Similarity 76.9%; Pred. No. 32;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASREAKKQVEAL 13  
 DB 171 ASREAKKKVEADL 183

RESULT 24  
 S05568  
 IGA receptor precursor - Streptococcus pyogenes  
 C/Species: Streptococcus pyogenes  
 C/Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 09-Jul-2004  
 C/Accession: S05568  
 R/Fritsch, E.; Heden, L.O.; Lindahl, G.  
 Mol. Microbiol. 3, 1111-1119, 1989  
 A/Title: Extensive sequence homology between IGA receptor and M proteins in Streptococcus  
 A/Reference number: S05568; MUID:50113892; PMID:2691841  
 A/Accession: S05568  
 A/Molecule type: DNA  
 A/Residues: 1-386 <PRI>  
 A/Cross-references: UNIPROT:P13050; EMBL:X15198; NID:G47350; PIDN:CAA33269.1; PID:G73631  
 C/Genetics:  
 A/Gene: artp4  
 C/Superfamily: M5 protein  
 C/Keywords: immunoglobulin receptor; transmembrane protein  
 F:1-41/Domain: signal sequence #status predicted <SIG>  
 F:42-386/Product: IGA receptor #status predicted <MAT>

Query Match 67.2%; Score 43; DB 2; Length 386;  
 Best Local Similarity 76.9%; Pred. No. 32;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASREAKKQVEAL 13  
 DB 215 ASREAKKKVEADL 227

RESULT 25  
 S54858  
 M protein precursor - Streptococcus pyogenes  
 C/Species: Streptococcus pyogenes  
 A/Variety: serotype POTTER C  
 C/Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
 C/Accession: S54858; S60822  
 R/Podbielski, A.  
 submitted to the EMBL Data Library, July 1993  
 A/Reference number: S54858  
 A/Accession: S54858  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-386 <POD>  
 A/Cross-references: UNIPROT:Q54829; EMBL:X74138; NID:G840899; PIDN:CAA52335.1; PID:G8409  
 R/Matmore, A.M.; Kapur, V.; Sullivan, D.J.; Mueser, J.M.; Kehoe, M.A.  
 Mol. Microbiol. 14, 619-631, 1994  
 A/Title: Non-congruent relationships between variation in emm gene sequences and the pop  
 A/Reference number: S60784; MUID:55198537; PMID:7891551  
 A/Accession: S60822  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 25-92 <HNA>  
 A/Cross-references: EMBL:U12000; NID:G533677; PIDN:AAA99616.1; PID:G1235844  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
 C/Superfamily: M5 protein

Query Match 67.2%; Score 43; DB 2; Length 386;  
 Best Local Similarity 76.9%; Pred. No. 32;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASREAKKQVEAL 13  
 DB 216 ASREAKKKVEADL 228

RESULT 26  
 A43715  
 M49 protein precursor - Streptococcus pyogenes  
 N/Alternate names: type 49 antiphagocytic M protein  
 C/Species: Streptococcus pyogenes  
 C/Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 09-Jul-2004  
 C/Accession: A43715; B42711; A54620; A92714; A60845; A28818  
 R/Hanes, E.J.; Cleary, P.P.  
 J. Bacteriol. 171, 6397-6408, 1989  
 A/Title: Identification of a divergent M protein gene and an M protein-related gene fami  
 A/Reference number: A43715; MUID:90078078; PMID:2687231  
 A/Accession: A43715  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-389 <HNA>  
 A/Cross-references: UNIPROT:P16947; GB:M23689; NID:G153696; PIDN:AAA26918.1; PID:G153697  
 R/Hanes, E.J.; Heath, D.G.; Cleary, P.P.  
 J. Bacteriol. 174, 4967-4976, 1992  
 A/Title: Architecture of the vir regions of group A streptococci parallels opacity fac  
 A/Reference number: A42711; MUID:5232431; PMID:1385809  
 A/Accession: B42711  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-82 <HNA>  
 A/Cross-references: GB:M8606; NID:G153630; PIDN:AAA26888.1; PID:G552004  
 A/Note: sequence extracted from NCBI backbone (NCBIN:108942, NCBI:P.108946)  
 R/Manjula, B.N.

Submitted to the Protein Sequence Database, August 1988  
 A:Reference number: A94620  
 A:Accession: A94620  
 A:Molecule type: protein  
 A:Residues: 46-47, 'K', 49-51, 'A', 53-188 <MAN>  
 A:Experimental source: type M49, strain B915  
 R:Khandke, K.M.; Fairwell, T.; Acharya, A.S.; Trus, B.L.; Manjula, B.N.  
 J. Biol. Chem. 263, 5075-5082, 1988  
 A:Title: Complete amino acid sequence of streptococcal PcpM49 protein, a nephritis-associated  
 A:Reference number: A92714; MUID:88186790; PMID:2451662  
 A:Accession: A92714  
 A:Molecule type: protein  
 A:Residues: 46-47, 'K', 49-51, 'A', 53-188 <KHA>  
 A:Experimental source: type M49, strain B915  
 R:Khandke, K.M.; Fairwell, T.; Manjula, B.N.  
 J. Exp. Med. 166, 151-162, 1987  
 A:Title: Difference in the structural features of streptococcal M proteins from nephritic  
 A:Reference number: A60845; MUID:87252908; PMID:3298523  
 A:Accession: A60845  
 A:Molecule type: protein  
 A:Residues: 46-47, 'K', 49-51, 'A', 53-105 <KH2>  
 A:Experimental source: type M49, strain B915  
 C:Genetics:  
 A:Gene: emm49  
 C:Superfamily: M5 protein  
 C:Keywords: coiled coil; dimer; transmembrane protein

Query Match 67.2%; Score 43; DB 2; Length 369;  
 Best Local Similarity 76.9%; Pred. No. 33;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKAL 13  
 |||||:|  
 Db 176 ASREAKKVEADL 188

RESULT 27  
 S37046  
 IGA receptor - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: S37046  
 R:Heden, L.; Lindahl, G.  
 submitted to the EMBL Data Library, May 1993  
 A:Description: Conserved and variable regions in protein Ayp, the IGA receptor of Streptococcus  
 A:Reference number: S37046  
 A:Accession: S37046  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-402 <HED>  
 A:Cross-references: UNIPROT:Q54876; EMBL:222751; NID:g3397628; PIDN:CAA80436.1; PID:g33976  
 C:Superfamily: M5 protein  
 C:Keywords: immunoglobulin receptor

Query Match 67.2%; Score 43; DB 2; Length 402;  
 Best Local Similarity 76.9%; Pred. No. 34;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKAL 13  
 |||||:|  
 Db 231 ASREAKKVEADL 243

RESULT 28  
 S23325  
 M2 protein precursor - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes  
 A:Variety: serotype M2  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C:Accession: S23325; S35761; S61078; S60785  
 R:Beesen, D.B.; Fischetti, V.A.  
 Infect. Immun. 60, 124-135, 1992  
 A:Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A strep

A:Reference number: S23325; MUID:92104662; PMID:1370269  
 A:Accession: S23325  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-407 <BES>  
 A:Cross-references: UNIPROT:P50468; EMBL:X61276; NID:g47369; PIDN:CAA43581.1; PID:g47370  
 R:Podbielski, A.  
 submitted to the EMBL Data Library, November 1992  
 A:Reference number: S35760  
 A:Accession: S35761  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-17 <POD>  
 A:Cross-references: EMBL:X69324  
 R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
 submitted to the EMBL Data Library, July 1994  
 A:Description: Noncongruent relationships between variation in emm1 gene sequences and the popl  
 A:Reference number: S61072  
 A:Accession: S61078  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 12-94 <WHA>  
 A:Cross-references: EMBL:U11958; NID:g533593; PIDN:AAA99574.1; PID:g1235820  
 R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
 Mol. Microbiol. 14, 619-631, 1994  
 A:Title: Non-congruent relationships between variation in emm gene sequences and the popl  
 A:Reference number: S60784; MUID:95198537; PMID:7891551  
 A:Accession: S60785  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 31-89 <WH2>  
 A:Cross-references: EMBL:U11958  
 C:Superfamily: M5 protein

Query Match 67.2%; Score 43; DB 2; Length 407;  
 Best Local Similarity 76.9%; Pred. No. 34;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKAL 13  
 |||||:|  
 Db 278 ASREAKKVEADL 290

RESULT 29  
 A46685  
 GTP-binding regulatory protein Gs alpha, form N1 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
 C:Accession: A46685  
 R:Crawford, J.A.; Mutchler, K.J.; Sullivan, B.E.; Lanigan, T.M.; Clark, M.S.; Russo, A.F.  
 J. Biol. Chem. 269, 9879-9885, 1993  
 A:Title: Neural expression of a novel alternatively spliced and polyadenylated Gsalpha tr  
 A:Reference number: A46685; MUID:9325868; PMID:8486667  
 A:Accession: A46685  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-91 <CRA>  
 A:Cross-references: UNIPROT:Q05087; GB:L10326; NID:g205609; PIDN:AAA41664.1; PID:g205610  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: nucleotide binding; P-loop  
 F:47-54/Region: nucleotide-binding motif A (P-loop)

Query Match 65.6%; Score 42; DB 2; Length 91;  
 Best Local Similarity 57.1%; Pred. No. 12;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKAL 14  
 |||||:|  
 Db 18 AQRANKIKKIKQLQ 31

RESULT 30  
 R0MSAI

GTP-binding regulatory protein Gs alpha-S1 chain (adenylate cyclase-stimulating) - mouse  
 N:Alternate names: guanine nucleotide binding protein Gs alpha-S1 chain; heterotrimeric  
 C:Species: Mus musculus (house mouse)  
 C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
 A:Accession: A25889  
 R:Sullivan, K.A.; Liao, Y.C.; Alborzi, A.; Beiderman, B.; Chang, F.H.; Masters, S.B.; Le  
 Proc. Natl. Acad. Sci. U.S.A. 83, 6687-6691, 1986  
 A:Title: Inhibitory and stimulatory G proteins of adenylate cyclase: cDNA and amino acid  
 A:Reference number: A94123; MUID:86313643; PMID:3092218  
 A:Accession: A25889  
 A:Molecule type: mRNA  
 A:Residues: 1-377 <STU>  
 A:Cross-references: UNIPROT:P04894; GB:M13964; NID:G193652; PID:AAA7745.1; PID:G309277  
 C:Keywords: alternative splicing; GTP binding; heterotrimer; nucleotide binding; P-loop;  
 C:Comment: The G protein is a family of guanine nucleotide-binding proteins that relay  
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
 race). It is specific for each type of G protein.  
 C:Comment: The Gs alpha chain is specific for G protein that is responsible for transduc  
 tic stimuli.  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: alternative splicing; GTP binding; heterotrimer; nucleotide binding; P-loop;  
 F:45-52/Region: nucleotide-binding motif A (P-loop)  
 F:275-278/Region: GTP-binding NKXD motif  
 F:51/Binding site: GTP (Gys) #status predicted  
 F:184/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 377;  
 Best Local Similarity 57.1%; Pred. No. 45;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKQVKEALE 14  
 Db 16 AQRANKKIEKQLQ 29  
 |||||::|||:  
 |||||::|||:

RESULT 31  
 RGHNA1  
 GTP-binding regulatory protein Gs alpha chain (adenylate cyclase-stimulating), splice fo  
 N:Alternate names: guanine nucleotide binding protein Gs alpha-S1 chain; heterotrimeric  
 C:Species: Homo sapiens (man)  
 C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 09-Jul-2004  
 A:Accession: C31927; B24366  
 R:Kozasa, T.; Itoh, H.; Tsukamoto, T.; Kaziro, Y.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 2081-2085, 1988  
 A:Title: Isolation and characterization of the human G-e-alpha gene.  
 A:Reference number: A31927; MUID:88176690; PMID:3127824  
 A:Accession: C31927  
 A:Molecule type: DNA  
 A:Residues: 1-71, 73-380 <KOZ3>  
 A:Cross-references: UNIPROT:P04895; UNIPROT:Q14433; GB:M21142; DDBJ:J03647; NID:G183402;  
 A:Note: splice form 3  
 A:Accession: D31927  
 A:Molecule type: DNA  
 A:Residues: 1-380 <KOZ4>  
 A:Cross-references: GB:M21142; DDBJ:J03647; NID:G183402; PID:AAA53149.1; PID:G386746  
 A:Note: splice form 4  
 R:Matteira, R.; Codina, J.; Croizat, A.; Kidd, V.; Woo, S.L.C.; Birnbaumer, L.  
 FEBS Lett. 206, 36-42, 1986  
 A:Title: Identification by molecular cloning of two forms of the alpha-subunit of the hu  
 A:Reference number: A24366; MUID:87005246; PMID:3093273  
 A:Accession: B24366  
 A:Molecule type: mRNA  
 A:Residues: 1-380 <MAT>  
 A:Cross-references: EMBL:X04409; NID:G31912; PID:CAA27997.1; PID:G31913  
 C:Comment: The G protein is a family of guanine nucleotide-binding proteins that relay  
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
 race). It is specific for each type of G protein.  
 C:Comment: The Gs alpha chain is specific for G protein that is responsible for transduc  
 tic stimuli.  
 C:Comment: See also PIR:RGHNA2.  
 C:Genetic: GDB:GNAS1; GNAS  
 A:Gene: GDB:GNAS1; GNAS  
 A:Cross-references: GDB:120628; OMIM:139320

A:Map position: 20q13.2-20q13.3  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprot  
 F:2-380/Product: GTP-binding regulatory protein Gs alpha chain, splice form 4 #status pr  
 F:2-11, 73-380/Product: GTP-binding regulatory protein Gs alpha chain, splice form 3 #sta  
 F:47-54/Region: nucleotide-binding motif A (P-loop)  
 F:276-281/Region: GTP-binding NKXD motif  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:3/Binding site: palmitate (Cys) (covalent) #status predicted  
 F:187/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 380;  
 Best Local Similarity 57.1%; Pred. No. 45;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKQVKEALE 14  
 Db 18 AQRANKKIEKQLQ 31  
 |||||::|||:  
 |||||::|||:

RESULT 32  
 RGHNAE  
 GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - Chine  
 N:Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G  
 C:Species: Cricetus griseus (Chinese hamster)  
 C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 19-Jan-2001  
 A:Accession: S08140  
 R:Mercken, L.; Moras, V.; Tocque, B.; Mayaux, J.F.  
 Nucleic Acids Res. 18, 662, 1990  
 A:Title: The cDNA sequence of the alpha-subunit of the Chinese hamster adenylate cyclase  
 A:Reference number: S08140; MUID:90175000; PMID:2106672  
 A:Accession: S08140  
 A:Molecule type: mRNA  
 A:Residues: 1-394 <MER>  
 A:Cross-references: EMBL:X17481; NID:G49491; PID:CAA3516.1; PID:G49492  
 C:Comment: The G protein is a family of guanine nucleotide-binding proteins that relay  
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
 race). It is specific for each type of G protein.  
 C:Comment: The Gs alpha chain is specific for G protein that is responsible for transduc  
 tic stimuli.  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprot  
 F:2-394/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>  
 F:47-54/Region: nucleotide-binding motif A (P-loop)  
 F:292-295/Region: GTP-binding NKXD motif  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:3/Binding site: palmitate (Cys) (covalent) #status predicted  
 F:201/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 394;  
 Best Local Similarity 57.1%; Pred. No. 47;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKQVKEALE 14  
 Db 18 AQRANKKIEKQLQ 31  
 |||||::|||:  
 |||||::|||:

RESULT 33  
 RGMSE2  
 GTP-binding regulatory protein Gs alpha-S2 chain (adenylate cyclase-stimulating) - mouse  
 N:Alternate names: guanine nucleotide binding protein Gs alpha-S2 chain; heterotrimeric  
 C:Species: Mus musculus (house mouse)  
 C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 19-Jan-2001  
 A:Accession: S03075  
 R:Rall, T.; Harrie, B.A.  
 FEBS Lett. 224, 365-371, 1987  
 A:Title: Identification of the lesion in the stimulatory GTP-binding protein of the unco  
 A:Reference number: S03075; MUID:88083563; PMID:2826231  
 A:Accession: S03075  
 A:Molecule type: mRNA  
 A:Residues: 1-394 <RAL>  
 A:Cross-references: GB:Y00703; NID:G51127; PID:CAA68695.1; PID:G51128

C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay a signal. The beta and gamma chains, required for GTPase activity, appear to be common to all three. It is specific for each type of G protein.

C;Comment: The Gs alpha chain is specific for G protein that is responsible for transducing cyclic stimuli.

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprotein

F;2-394/Product: GTP-binding regulatory protein Gs alpha-S2 chain #status predicted <MAT>

F;47-54/Region: nucleotide-binding motif A (P-loop)

F;292-295/Region: GTP-binding NKXD motif

F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F;3/Binding site: palmitate (Cys) (covalent) #status predicted

F;201/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 394;  
Best Local Similarity 57.1%; Pred. No. 47;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14  
Db 18 AOREANKKIKERQLO 31

RESULT 34  
RGHYA2  
GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - golden hamster

N;Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G protein

C;Species: Mesocricetus auratus (golden hamster)

C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004

C;Accession: S10508

R;Conner, D.A.; Feldman, A.M.; Van Dop, C.

Nucleic Acids Res. 18, 4279, 1990

A;Title: cDNA sequence for the alpha subunit of the guanine nucleotide-binding protein t

A;Reference number: S10508; MUID:90332451; PMID:2115997

A;Accession: S10508

A;Molecule type: mRNA

A;Residues: 1-394 <CON>

C;Cross-references: UNIPROT:P04894; EMBL:X53139; NID:9496642; PIDN:CAA78161.1; PID:949643

C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay a signal. The beta and gamma chains, required for GTPase activity, appear to be common to all three. It is specific for each type of G protein.

C;Comment: The Gs alpha chain is specific for G protein that is responsible for transducing cyclic stimuli.

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprotein

F;2-394/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>

F;47-54/Region: nucleotide-binding motif A (P-loop)

F;292-295/Region: GTP-binding NKXD motif

F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F;3/Binding site: palmitate (Cys) (covalent) #status predicted

F;201/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 394;  
Best Local Similarity 57.1%; Pred. No. 47;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14  
Db 18 AOREANKKIKERQLO 31

RESULT 35  
RGRTA2  
GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - rat

N;Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G protein

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004

C;Accession: A27423; C24882

R;Jones, D.T.; Reed, R.R.

J. Biol. Chem. 262, 14241-14249, 1987

A;Title: Molecular cloning of five GTP-binding protein cDNA species from rat olfactory r

A;Reference number: A27423; MUID:88007678; PMID:2820999

A;Accession: A27423

A;Molecule type: mRNA

A;Residues: 1-394 <CON>

C;Cross-references: UNIPROT:P04894; GB:M17525; NID:9203171; PIDN:AAA40827.1; PID:9203172

R;Itoh, H.; Kozasa, T.; Negata, S.; Nakamura, S.; Katada, T.; Ui, M.; Iwai, S.; Ohtsuka, Proc. Natl. Acad. Sci. U.S.A. 83, 3776-3780, 1986

A;Title: Molecular cloning and sequence determination of cDNAs for alpha subunits of the

A;Reference number: A94707; MUID:86233317; PMID:3086867

A;Accession: C24882

A;Molecule type: mRNA

A;Residues: 1-394 <ITO>

A;Cross-references: GB:M12673; NID:9204441; PIDN:AAA41261.1; PID:9204442

C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay a signal. The beta and gamma chains, required for GTPase activity, appear to be common to all three. It is specific for each type of G protein.

C;Comment: The Gs alpha chain is specific for G protein that is responsible for transducing cyclic stimuli.

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: blocked amino end; GTP binding; lipoprotein; myristylation; nucleotide binding

F;2-394/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>

F;47-54/Region: nucleotide-binding motif A (P-loop)

F;292-295/Region: GTP-binding NKXD motif

F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F;3/Binding site: palmitate (Cys) (covalent) #status predicted

F;201/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 394;  
Best Local Similarity 57.1%; Pred. No. 47;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14  
Db 18 AOREANKKIKERQLO 31

RESULT 36  
S33458  
GTP-binding regulatory protein Gs alpha chain - dog

C;Species: Canis lupus familiaris (dog)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S33458

R;Ishikawa, Y.; Homcy, C.J.

submitted to the EMBL Data Library, June 1992

A;Description: cDNA sequence for alpha subunit of stimulatory guanine nucleotide binding

A;Reference number: S33458

A;Accession: S33458

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-394 <ISH>

C;Cross-references: UNIPROT:P04895; EMBL:Z12168; NID:9311336; PIDN:CAA78161.1; PID:931133

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: GTP binding; nucleotide binding; P-loop

F;47-54/Region: nucleotide-binding motif A (P-loop)

F;292-295/Region: GTP-binding NKXD motif

Query Match 65.6%; Score 42; DB 2; Length 394;  
Best Local Similarity 57.1%; Pred. No. 47;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14  
Db 18 AOREANKKIKERQLO 31

RESULT 37  
RGHND2  
GTP-binding regulatory protein Gs alpha chain (adenylate cyclase-stimulating), splice form 1

N;Alternate names: guanine nucleotide binding protein Gs alpha-S2 chain; heterotrimeric G protein

N;Contains: GTP-binding regulatory protein Gs alpha chain, splice form 1

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1990 #sequence\_revision 15-Jun-1996 #text\_change 09-Jul-2004

C;Accession: B31927; A31927; A24366; S02122; A25919

R;Kozasa, T.; Itoh, H.; Tsukamoto, T.; Kariro, Y.

Proc. Natl. Acad. Sci. U.S.A. 85, 2081-2085, 1988

A/Title: Isolation and characterization of the human G-s-alpha gene.  
 A/Reference number: A31927; MUID:88176890; PMID:3127824  
 A/Accession: B31927  
 A/Molecule type: DNA  
 A/Residues: 1-395 <KOZ2>  
 A/Cross-references: UNIPROT:Q96H70; GB:M21142; DDBJ:J03647; NID:g183402; PIDN:AAA53147.1  
 A/Note: splice form 2  
 A/Accession: A31927  
 A/Molecule type: DNA  
 A/Residues: 1-86,88-395 <KOZ1>  
 A/Cross-references: GB:M21142; DDBJ:J03647; NID:g183402; PIDN:AAA53146.1; PID:g386743  
 A/Note: splice form 1  
 A/Molecule type: DNA  
 A/Residues: 1-86,88-395 <MAT>  
 A/Cross-references: EMBL:X04408; NID:g31914; PIDN:CAA27996.1; PID:g31915  
 A/Note: R.Harris, B.A.  
 A/Title: Complete cDNA sequence of a human stimulatory GTP-binding protein alpha subunit  
 A/Reference number: S02122; MUID:88233954; PMID:3131741  
 A/Accession: S02122  
 A/Molecule type: mRNA  
 A/Residues: 1-5,'T',7-86,88-395 <HAR>  
 A/Cross-references: EMBL:X07036; NID:g31951; PIDN:CAA30084.1; PID:g31952  
 A/Note: R.Brady, P.J. Carter, A.J. Simons, C.J. Guo, V.J. Puckett, C.J. Kamholz, J.J. Spiegel, A.J. Nire  
 A/Title: Human cDNA clones for four species of G-alpha-s-signal transduction protein.  
 A/Reference number: A25919; MUID:87067419; PMID:3024154  
 A/Accession: A25919  
 A/Molecule type: mRNA  
 A/Residues: 12-230,'Q',232-395 <BRA>  
 A/Cross-references: GB:M14631; NID:g183416; PIDN:AAA52583.1; PID:g386748  
 A/Note: The authors translated the codon CAA for residue 221 as Glu  
 A/Comment: The G protein are a family of guanine nucleotide-binding proteins that relay  
 a signal. The beta and gamma chains, required for GTPase activity, appear to be common to all  
 classes. It is specific for each type of G protein.  
 A/Comment: The Gs alpha chain is specific for G protein that is responsible for transduc-  
 toric stimulation.  
 A/Comment: See also PIR:RGHUA1.  
 A/Accession: C/Genetics:  
 A/Accession: GDB:120628; OMIM:139320  
 A/Map position: 20q13.2-20q13.3  
 A/Introns: 47/1, 71/2, 86/2, 105/3, 145/3, 178/2, 196/3, 221/2, 241/1, 281/2, 325/1, 347  
 A/Supernote: GTP-binding regulatory protein Gs alpha chain  
 A/Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprot  
 A/2-395/Product: GTP-binding regulatory protein Gs alpha chain, splice form 2 #status pr  
 A/2-86,88-395/Product: GTP-binding regulatory protein Gs alpha chain, splice form 1 #sta  
 A/47-54/Region: nucleotide-binding motif A (P-loop)  
 A/232-266/Region: GTP-binding NKXD motif  
 A/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 A/3/Binding site: palmitate (Cys) (covalent) #status predicted  
 A/202/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 395;  
 Best Local Similarity 57.1%; Pred. No. 47;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 18 AGRANKKIKKQLO 31

RESULT 38  
 RGP6A2  
 GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - pig  
 N/Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G  
 C/Species: Sus scrofa domestica (domestic pig)  
 C/Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004

C/Accession: S18963  
 R/Roth, D.A.; Kay, R.A.M.; Hammond, H.K.  
 submitted to the EMBL Data Library, January 1992  
 A/Reference number: S18963  
 A/Accession: S18963  
 A/Molecule type: mRNA  
 A/Residues: 1-397 <ROT>  
 A/Cross-references: UNIPROT:P29797; EMBL:X63893; NID:g1957; PIDN:CAA45355.1; PID:g1958  
 A/Comment: The G protein are a family of guanine nucleotide-binding proteins that relay  
 a signal. The beta and gamma chains, required for GTPase activity, appear to be common to all  
 classes. It is specific for each type of G protein.  
 A/Comment: The Gs alpha chain is specific for G protein that is responsible for transduc-  
 toric stimulation.  
 A/Supernote: GTP-binding regulatory protein Gs alpha chain  
 A/Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprot  
 A/2-397/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>  
 A/47-54/Region: nucleotide-binding motif A (P-loop)  
 A/277-280/Region: GTP-binding NKXD motif  
 A/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 A/3/Binding site: palmitate (Cys) (covalent) #status predicted  
 A/186/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 397;  
 Best Local Similarity 57.1%; Pred. No. 47;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
 |||||  
 Db 18 AGRANKKIKKQLO 31

RESULT 39  
 T22180  
 hypothetical protein F44F1.6a - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 A/Accession: T22180  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-493 <WIL>  
 A/Cross-references: UNIPROT:Q9J3F0; EMBL:Z81083; PIDN:CAB54248.1; GSPDB:GN00019; CESP:F4  
 A/Experimental source: clone F44F1  
 A/Accession: C/Genetics:  
 A/Accession: A/Map position: 1  
 A/Map position: 1

Query Match 65.6%; Score 42; DB 2; Length 493;  
 Best Local Similarity 69.2%; Pred. No. 58;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKAL 13  
 |||||  
 Db 474 ASREAKKQYENIM 486

RESULT 40  
 T06930  
 H+transporting two-sector ATPase (BC 3.6.3.14) epsilon chain - Cyanophora paradoxa cyan  
 C/Species: cyanella Cyanophora paradoxa  
 C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
 A/Accession: T06930  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Description: Nucleotide sequence of the cyanella genome from Cyanophora paradoxa.  
 A/Reference number: Z15840  
 A/Accession: T06930  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-132 <STI>

A/Cross-references: UNIPROT:P48083; EMBL:U30821; NID:g1016083; PIDN:AAA81273.1; PID:g101  
A/Experimental source: strain Pringsheim LB555  
C/Genetics:  
A/Gene: atpE  
A/Genome: Cyanelle  
C/Superfamily: H<sup>+</sup>-transporting ATP synthase epsilon chain  
C/keywords: ATP biosynthesis; cyanelle; hydrolase; membrane-associated complex; thylakoi

Query Match 64.1%; Score 41; DB 2; Length 132;

Best Local Similarity 90.0%; Pred. No. 24;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AKQVEKALE 14  
||| |||||  
Db 92 AKQVEKALE 101

Search completed: June 13, 2005, 20:53:36  
Job time : 16.6279 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 20:30:50 ; Search time 74.2126 Seconds  
(without alignments)  
96.576 Million cell updates/sec

Title: US-10-706-275-1  
Perfect score: 64  
Sequence: 1 ASREAKKOVERALE 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	100	2	Q54639 streptococc
2	64	100.0	100	2	Q54640 streptococc
3	64	100.0	100	2	Q9R3A1 streptococc
4	64	100.0	198	2	Q54832 streptococc
5	64	100.0	208	2	P95824 streptococc
6	64	100.0	208	2	P95825 streptococc
7	64	100.0	208	2	P95826 streptococc
8	64	100.0	212	2	Q84DD3 streptococc
9	64	100.0	237	2	Q6TFR5 streptococc
10	64	100.0	251	2	Q6V9Q3 streptococc
11	64	100.0	279	2	Q8G187 streptococc
12	64	100.0	282	2	Q8G1A6 streptococc
13	64	100.0	303	2	Q8G198 streptococc
14	64	100.0	307	2	Q8G184 streptococc
15	64	100.0	314	2	Q8G1B0 streptococc
16	64	100.0	317	2	Q8G192 streptococc
17	64	100.0	319	2	Q8G1A2 streptococc
18	64	100.0	322	2	Q8G1A4 streptococc
19	64	100.0	326	2	Q8G191 streptococc
20	64	100.0	340	2	Q8G1A1 streptococc
21	64	100.0	345	2	Q8G193 streptococc
22	64	100.0	347	2	Q93R06 streptococc
23	64	100.0	362	2	Q8G189 streptococc
24	64	100.0	362	2	Q8G190 streptococc
25	64	100.0	388	1	PAM_STRPY streptococc
26	64	100.0	400	2	Q8N279 streptococc
27	64	100.0	408	2	Q54837 streptococc
28	64	100.0	435	2	Q9AMM3 streptococc
29	64	100.0	436	2	Q54839 streptococc
30	64	100.0	441	2	Q55246 streptococc
31	64	100.0	443	2	Q54703 streptococc

32	64	100.0	454	2	Q55278 streptococc
33	64	100.0	454	2	Q840T7 streptococc
34	64	100.0	457	2	Q54510 streptococc
35	64	100.0	465	2	Q83XW0 streptococc
36	64	100.0	471	2	Q93SL9 streptococc
37	64	100.0	472	2	Q55279 streptococc
38	64	100.0	475	2	Q33631 streptococc
39	64	100.0	483	1	M6_STRPY streptococc
40	64	100.0	484	2	Q05464 streptococc
41	64	100.0	484	2	Q10372 streptococc
42	64	100.0	484	2	Q99XV0 streptococc
43	64	100.0	488	2	Q54830 streptococc
44	64	100.0	492	1	M5_STRPS streptococc
45	64	100.0	500	2	Q9RHV2 streptococc

## ALIGNMENTS

```

RESULT 1
Q54639 PRELIMINARY; PRT; 100 AA.
AC Q54639;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M1 protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95172752; PubMed=7868273;
RA Musser J.M., Kapur V., Szeto J., Pan X., Swanson D.S., Martin D.R.;
RT "Genetic diversity and relationships among Streptococcus pyogenes
RT strains expressing serotype M1 protein: recent intercontinental spread
RT of a subclone causing episodes of invasive disease.";
RL Infect. Immun. 63:994-1003(1995).
DR EMBL; U20103; AA85116.1; -.
DR HSSP; P13276; 1BQ1.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1 1
FT NON_TER 100 100
SQ SEQUENCE 100 AA; 11243 MW; 9775831FA25DC463 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERALE 14
Db 50 ASREAKKOVERALE 63

RESULT 2
Q54640 PRELIMINARY; PRT; 100 AA.
AC Q54640;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M1 protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95172752; PubMed=7868273;
RA Musser J.M., Kapur V., Szeto J., Pan X., Swanson D.S., Martin D.R.;

```

```
RT "Genetic diversity and relationships among Streptococcus pyogenes
RT strains expressing serotype M1 protein: recent intercontinental spread
RT of a subclone causing epistaxis M1 protein: recent intercontinental spread
RL Infect. Immun. 63:994-1003(1995).
DR EMBL; U20104; AAA85117.1; -.
DR HSSP; P13276; IEQ1.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 100 AA; 11287 MW; 9773331914EDC2D3 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14
Db 50 ASREAKKQVEKALE 63

RESULT 3
Oy 09R3A1 PRELIMINARY; PRT; 100 AA.
ID 09R3A1
AC 09R3A1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE M1 protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95172752; PubMed=7668273;
RA Musser J.M., Kapur V., Szeto J., Pan X., Swanson D.S., Martin D.R.;
RT "Genetic diversity and relationships among Streptococcus pyogenes
RT strains expressing serotype M1 protein: recent intercontinental spread
RT of a subclone causing epistaxis M1 protein: recent intercontinental spread
RL Infect. Immun. 63:994-1003(1995).
DR EMBL; U20102; AAA85115.1; -.
DR HSSP; P13276; IEQ1.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 100 AA; 11345 MW; 9773331C00EDC2D3 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14
Db 50 ASREAKKQVEKALE 63

RESULT 4
Oy 054832 PRELIMINARY; PRT; 198 AA.
ID 054832
AC 054832;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm3;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
FT NON_TER 1
FT NON_TER 1
```

```
Ox NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3-3/317;
RX MEDLINE=93062420; PubMed=1435517;
RA Podbielski A., Baid R., Kautfohl A.;
RT "The group A streptococcal M-type 3 protein gene exhibits a C terminus
RT typical for class I M proteins."
RL Med. Microbiol. Immunol. 181:209-213 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3-3/317;
RA Podbielski A., Kautfohl A.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (by similarity).
DR EMBL; X66816; CAA47295.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M; 1.
DR PRINTS; PR00015; GP05ANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 198 AA; 21550 MW; A7388880947155D5 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14
Db 49 ASREAKKQVEKALE 62

RESULT 5
Oy P95824 PRELIMINARY; PRT; 208 AA.
ID P95824
AC P95824;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE M protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=M nontypeable group A;
RA Brandt E.R., Good M.F.;
RT Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (by similarity).
DR EMBL; U65899; AAB40640.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M; 1.
DR PRINTS; PR00015; GP05ANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
FT NON_TER 1
FT NON_TER 1
```



SO SEQUENCE 208 AA; 22695 MW; 013FBBAB0A97FF42 CRC64;

Query Match Best Local Similarity 100.0%; Score 64; DB 2; Length 208;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERKALE 14  
Db 59 ASREAKKOVERKALE 72

## RESULT 6

ID P95825 PRELIMINARY; PRT; 208 AA.  
AC P95825;  
DT 01-MAY-1997 (TReMBLrel. 03, Created)  
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
DE M protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=reference;  
RA Brandt E.R., Good M.F.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).

DR EMBL; U65900; AAB40641.1; -;  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 1.  
DR PRINTS; PR00015; GP0SANCHOR.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KM Cell wall; Peptidoglycan-anchor.  
FT NON\_TER 1  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 22817 MW; 790342752FB17720 CRC64;

Query Match Best Local Similarity 100.0%; Score 64; DB 2; Length 208;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERKALE 14  
Db 59 ASREAKKOVERKALE 72

## RESULT 7

ID P95826 PRELIMINARY; PRT; 208 AA.  
AC P95826;  
DT 01-MAY-1997 (TReMBLrel. 03, Created)  
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RHD152-;  
RA Brandt E.R., Good M.F.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by

CC an amide bond (By similarity).

DR EMBL; U66005; AAB40642.1; -;  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF02370; M; 1.

DR PRINTS; PR00015; GP0SANCHOR.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KM Cell wall; Peptidoglycan-anchor.  
FT NON\_TER 1  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 22565 MW; 79972A987324729B CRC64;

Query Match Best Local Similarity 100.0%; Score 64; DB 2; Length 208;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERKALE 14  
Db 59 ASREAKKOVERKALE 72

## RESULT 8

ID Q84DD3 PRELIMINARY; PRT; 212 AA.  
AC Q84DD3;  
DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Beall B.W.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY225412; AAO67526.1; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.

FT NON\_TER 1  
FT CHAIN 23  
FT NON\_TER 212  
SQ SEQUENCE 212 AA; 24116 MW; 0A7B56F0FCAEF26 CRC64;

Query Match Best Local Similarity 100.0%; Score 64; DB 2; Length 212;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASREAKKOVERKALE 14  
Db 197 ASREAKKOVERKALE 210

## RESULT 9

ID Q6TLR5 PRELIMINARY; PRT; 237 AA.  
AC Q6TLR5;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.

```

OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS241;
RX MEDLINE=22894607; PubMed=14532198;
RA Moses A.E., Hidalgo-Grass C., Dan-Goor M., Jaffe J., Shetzigovsky I.,
RA Ravin M., Koreman Z., Cohen-Poradosu R., Nir-Paz R.;
RT "emm typing of M nontypeable invasive group A streptococcal isolates
  in Israel";
RL J Clin Microbiol. 41:4655-4659(2003).
DR EMBL; AY394538; AA094530.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1 1
SQ SEQUENCE 237 AA; 27027 MW; 47CF315DD4EB5F2 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKALE 14
Db 192 ASREAKQVEKALE 205

RESULT 10
Q8V903 PRELIMINARY; PRT; 251 AA.
AC Q8V903;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=J58;
RC Moses A.E., Hidalgo-Grass C., Dan-Goor M., Jaffe J., Shetzigovsky I.,
RA Ravin M., Koreman Z., Cohen-Poradosu R., Nir-Paz R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY346386; AAQ73206.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28938 MW; 2A6602AAA637D11 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 0.098;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKALE 14
Db 221 ASREAKQVEKALE 234

RESULT 11
Q8GL87 PRELIMINARY; PRT; 279 AA.
AC Q8GL87;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CX NCBI_TaxID=1314;
```

```

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dvalli-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,
RA McMillan D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY394420; AAN64693.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 2.
FT NON_TER 1 1
FT NON_TER 279 279
SQ SEQUENCE 279 AA; 31224 MW; 16A600455BC3A0 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKALE 14
Db 224 ASREAKQVEKALE 237

RESULT 12
Q8GLA6 PRELIMINARY; PRT; 282 AA.
ID Q8GLA6
AC Q8GLA6;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dvalli-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,
RA McMillan D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY394401; AAN64674.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1 1
FT NON_TER 282 282
SQ SEQUENCE 282 AA; 32157 MW; 5C799B0AA4323541 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKQVEKALE 14
Db 226 ASREAKQVEKALE 239

RESULT 13
Q8GL98 PRELIMINARY; PRT; 303 AA.
ID Q8GL98
AC Q8GL98;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CX NCBI_TaxID=1314;
```

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dyll-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,  
 RA McMillan D.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY139409; AAN64682.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF02370; M; 2.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 303 AA; 34562 MW; F76F37540E16CD1B CRC64;  
 Query Match 100.0%; Score 64; DB 2; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVERKALE 14  
 Db 246 ASREAKKOVERKALE 259  
 RESULT 14  
 ID Q8GL84 PRELIMINARY; PRT; 307 AA.  
 AC Q8GL84;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE M protein (Fragment).  
 GN Name=emm;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dyll-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,  
 RA McMillan D.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY139433; AAN64696.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF02370; M; 2.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 307 AA; 34955 MW; 226822938B66E0E CRC64;

Query Match 100.0%; Score 64; DB 2; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVERKALE 14  
 Db 251 ASREAKKOVERKALE 264  
 RESULT 15  
 ID Q8GLB0 PRELIMINARY; PRT; 314 AA.  
 AC Q8GLB0;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE M protein (Fragment).  
 GN Name=emm;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dyll-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,

RA McMillan D.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY139397; AAN64670.1; -.  
 DR HSSP; P04268; IIC2.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF02370; M; 2.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 314 AA; 35613 MW; 66173BEB74C9EFC CRC64;

Query Match 100.0%; Score 64; DB 2; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVERKALE 14  
 Db 259 ASREAKKOVERKALE 272  
 RESULT 16  
 ID Q8GL92 PRELIMINARY; PRT; 317 AA.  
 AC Q8GL92;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE M protein (Fragment).  
 GN Name=emm;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dyll-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,  
 RA McMillan D.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY139415; AAN64688.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF02370; M; 2.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 317 AA; 36056 MW; F2026105F02D4888 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKOVERKALE 14  
 Db 262 ASREAKKOVERKALE 275  
 RESULT 17  
 ID Q8GLA2 PRELIMINARY; PRT; 319 AA.  
 AC Q8GLA2;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE M protein (Fragment).  
 GN Name=emm;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dyll-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,  
 RA McMillan D.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY139405; AAN64678.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 319  
SQ SEQUENCE 319 AA; 36283 MW; 1797D2026105F02D CRC64;

Query Match 100.0%; Score 64; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 262 ASREAKKOVERKALE 275

## RESULT 18

O8GLA4 PRELIMINARY; PRT; 322 AA.  
AC O8GLA4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dyall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139403; AAN64676.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 322  
SQ SEQUENCE 322 AA; 36990 MW; F179D8DF126CED92 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 322;  
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 265 ASREAKKOVERKALE 278

## RESULT 19

O8GL91 PRELIMINARY; PRT; 326 AA.  
AC O8GL91;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dyall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139416; AAN64689.1; -.  
DR HSP; P02647; IAVI.  
DR GO; GO:0016020; C:membrane; IEA.

DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 326  
SQ SEQUENCE 326 AA; 37290 MW; 7186FE5FEB970C7 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 326;  
Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 271 ASREAKKOVERKALE 284

## RESULT 20

O8GLA1 PRELIMINARY; PRT; 340 AA.  
AC O8GLA1;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dyall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139406; AAN64679.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 340  
SQ SEQUENCE 340 AA; 38495 MW; 2222F0DE240268CD CRC64;

Query Match 100.0%; Score 64; DB 2; Length 340;  
Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERKALE 14  
Db 283 ASREAKKOVERKALE 296

## RESULT 21

O8GL93 PRELIMINARY; PRT; 345 AA.  
AC O8GL93;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dyall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139414; AAN64687.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.

DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 1  
SEQUENCE 345 AA; 39030 MW; 7BAFCFBFA82F0C6 CRC64;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 64; DB 2; Length 345;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 290 ASREAKKQVEKALE 303

RESULT 22  
ID Q93R06 PRELIMINARY; PRT; 347 AA.  
AC Q93R06; 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gowan S.N., Martin D.R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ319534; CAC39213.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 347  
SQ SEQUENCE 347 AA; 39229 MW; 97352E7278D83BD CRC64;

Query Match 100.0%; Score 64; DB 2; Length 347;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 292 ASREAKKQVEKALE 305

RESULT 23  
ID Q8GL89 PRELIMINARY; PRT; 362 AA.  
AC Q8GL89; 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dvall-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139418; AAN64691.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 362  
SQ SEQUENCE 362 AA; 41017 MW; 685681D5E2FAC770 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 307 ASREAKKQVEKALE 320

RESULT 24  
ID Q8GL90 PRELIMINARY; PRT; 369 AA.  
AC Q8GL90; 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dvall-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139417; AAN64690.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
FT NON\_TER 369  
SQ SEQUENCE 369 AA; 41998 MW; EA81172C872012BD CRC64;

Query Match 100.0%; Score 64; DB 2; Length 369;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREAKKQVEKALE 14  
Db 314 ASREAKKQVEKALE 327

RESULT 25  
ID PAM\_STRPY STANDARD; PRT; 388 AA.  
AC P49054; 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DE 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Plasmidogen-binding group A streptococcal M-like protein PAM precursor (Fragment).  
GN Name=pam; Synonyms=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-41.  
RC STRAIN=AP53 / Serotype M53;  
RX MEDLINE=94064605; PubMed=8244975;  
RA Berge A., Sjoebing U.;  
RT "PAM, a novel plasmidogen-binding protein from Streptococcus pyogenes";  
RL J. Biol. Chem. 268:25417-25424(1993).  
RN [2]  
RP CHARACTERIZATION OF PLASMINOGEN BINDING AND MUTAGENESIS.  
RC STRAIN=AP53 / Serotype M53;  
RX PubMed=8748039;  
RA Carlsson Wistedt A., Ringdahl U., Mueller-Esterl W., Sjoebing U.;  
RT "Identification of a plasmidogen-binding motif in PAM, a bacterial

```

RT surface protein.";
RL Mol. Microbiol. 18:569-578(1995).
CC -1- FUNCTION: Binds to human plasminogen (and plasmin) via its kringle
CC repeats. Also binds to albumin, immunoglobulin G and fibrinogen.
CC Could provide the bacteria with a mechanism for invasion, as
CC streptococcal-bound plasmin could permit tissue penetration.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -1- MISCELLANEOUS: PAM has more than one binding site for plasminogen;
CC it is thought that each of the A-repeats can bind one plasminogen
CC molecule.
CC -1- SIMILARITY: Belongs to the M protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22219; CA80222.1; -.
DR PIR; A49545; A49545.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M_2.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; PARTIAL.
DR Cell wall; Direct protein sequencing; Peptidoglycan-anchor; Repeat;
KW Signal.
FT NON_TER <1 29 1
FT SIGNAL 30 384
FT CHAIN
FT PROPEP 385 >388 Potential.
FT DOMAIN 85 113 Plasminogen-binding group A streptococcal
FT DOMAIN 91 116 M-like protein PAM. (Potential).
FT REPEAT 91 103 Able to bind plasminogen.
FT REPEAT 104 103 2 X approximate tandem repeats, type a.
FT REPEAT 104 103 A-1.
FT REPEAT 147 161 A-2.
FT REPEAT 147 153 2 X tandem repeats, type b.
FT REPEAT 154 161 B-1.
FT REPEAT 154 161 B-2.
FT DOMAIN 163 204 3 X tandem repeats, type c.
FT REPEAT 163 204 C-1.
FT REPEAT 205 246 C-2.
FT REPEAT 247 278 C-3 (incomplete).
FT DOMAIN 344 380 Gly/Pro-rich.
FT SITE 381 385 LPTXG sorting signal (Potential).
FT MOD_RES 384 384 Pentaglycyl murein peptidoglycan amidated
FT threonine (potential).
FT MUTAGEN 82 82 K->A: No change in plasminogen binding.
FT MUTAGEN 98 98 K->A: 50-fold decrease in plasminogen
FT binding.
FT MUTAGEN 111 111 K->A: 2-fold decrease in plasminogen
FT binding.
FT NON_TER 388 388
FT SEQUENCE 388 AA; 43629 MW; EEEBC4FD962CCDB12 CRC64;
SQ
Query Match 100.0%; Score 64; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASREAKKQVEKALE 14
DB 266 ASREAKKQVEKALE 279

```

```

DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M18 protein.
GN Name=emm18; OrderedLocustName=spym18 2076;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WGA58232;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Bardhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Studevant D.E., Ricklefs S.M., Porcella S.F.,
RA Pakkins L.D., Bares S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
CC
CC EMBL; AB010111; AAL98546.1; -.
CC GO; GO:0009986; C:cell surface; IEA.
CC GO; GO:0005618; C:cell wall; IEA.
CC GO; GO:0016020; C:membrane; IEA.
CC InterPro; IPR011000; Apolp_III_like.
CC InterPro; IPR005877; Gpos_YsIRK.
CC InterPro; IPR001899; Gram_pos_anchor.
CC InterPro; IPR003345; M_repeat.
CC Pfam; PF00746; Gram_pos_anchor; 1.
CC Pfam; PF02370; M_1.
CC Pfam; PF04650; YsIRK_signal; 1.
CC PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 400 AA; 44645 MW; D73F2A1A8BC7813F CRC64;
QY 1 ASREAKKQVEKALE 14
DB 252 ASREAKKQVEKALE 265
Query Match 100.0%; Score 64; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASREAKKQVEKALE 14
DB 252 ASREAKKQVEKALE 265

```

```

RESULT 27 PRELIMINARY; PRT; 408 AA.
ID OS4837
AC OS4837;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M protein type 41.
GN Name=emm41;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C101/103/2;
RX MEDLINE=93204905; PubMed=8455563;
RA Podbielski A.;
RT "Three different types of organization of the vir region in group A
RT streptococci.";
RT Mol. Gen. Genet. 237:287-300(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).

```

DR EMBL, X58178; CAA41167.1; -.  
 DR PIR, S30283; S30283.  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR011000; Apolp\_TII\_like.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M\_2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LpYTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall, peptidoglycan-anchor.  
 SQ SEQUENCE 408 AA; 45630 MW; 5F449A7645B9696C CRC64;

Query Match 100.0%; Score 64; DB 2; Length 408;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERALE 14  
 |||||  
 Db 259 ASREAKKOVERALE 272

RESULT 28  
 Q9AMM3 PRELIMINARY; PRT; 435 AA.  
 ID Q9AMM3  
 AC Q9AMM3  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DR M S14547 protein.  
 GN Name=emmS14547;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 CX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Esahgh M., Ali M., Yusoff K., Jamal F.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 an amide bond (By similarity).  
 DR EMBL; AF324061; AKL1612.1; -.  
 DR HSBP; P02647; IAV1.  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M\_2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LpYTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall, peptidoglycan-anchor.  
 SQ SEQUENCE 435 AA; 49022 MW; A64C48F7137AE046 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERALE 14  
 |||||  
 Db 293 ASREAKKOVERALE 306

RESULT 29  
 O54839 PRELIMINARY; PRT; 436 AA.  
 ID O54839  
 AC O54839  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE M protein type 52.  
 GN Name=emm52;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 CX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=871/14/3;  
 RX MEDLINE=93204905; PubMed=8455563;  
 RA Podbielski A.;  
 RT "Three different types of organization of the vir regulon in group A  
 streptococci.";  
 RL Mol.Gen. Genet. 237:287-300(1993).  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 an amide bond (By similarity).  
 DR EMBL; X58179; CAA41168.1; -.  
 DR PIR; S30284; S30284.  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M\_2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LpYTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall, peptidoglycan-anchor.  
 SQ SEQUENCE 436 AA; 48951 MW; 4C5720F98F2DAE89 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKOVERALE 14  
 |||||  
 Db 287 ASREAKKOVERALE 300

RESULT 30  
 Q55246 PRELIMINARY; PRT; 441 AA.  
 ID Q55246  
 AC Q55246  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE M protein.  
 GN Name=emm1;  
 OS Streptococcus sp.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 CX NCBI\_TaxID=1306;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Group G;  
 RX MEDLINE=9522925; PubMed=7714192;  
 RA Schmitzler N., Podbielski A., Baumgarten G., Mignon M., Kaufhold A.;  
 RT "M or M-like protein gene polymorphisms in group G streptococci.";  
 RL J. Clin. Microbiol. 33:356-363(1995).  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 an amide bond (By similarity).

DR EMBL; X60098; CAA42694.1; -;  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR011000; Apolp\_III\_like.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M\_2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KM Cell wall; peptidoglycan-anchor.  
 SQ SEQUENCE 441 AA; 49868 MW; 33DF991E365D9455 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 0.17;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14  
 Db 292 ASREAKKQVEKALE 305

RESULT 31  
 ID 054703 PRELIMINARY; PRT; 443 AA.  
 AC 054703;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Emu18.1 protein.  
 GN Name=emu18.1;  
 OS Streptococcus pyogenes.  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=type 18;  
 RX MEDLINE=96201553; PubMed=8613352;  
 RA Dale J.B., Maaburn R.G., Marques M.B., Wessels M.R.;  
 RT "Hyaluronate capsule and surface M protein in resistance to  
 RT opsonization of group A streptococci";  
 RL Infect. Immun. 64:1495-1501(1996).  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (By similarity).  
 DR EMBL; U29585; AAB03086.1; -;  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR011000; Apolp\_III\_like.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M\_3.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KM Cell wall; peptidoglycan-anchor.  
 SQ SEQUENCE 443 AA; 49365 MW; 239B0E7F0ADB5EAB CRC64;

Query Match 100.0%; Score 64; DB 2; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 0.17;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14

Db 294 ASREAKKQVEKALE 307

RESULT 32  
 ID 055278 PRELIMINARY; PRT; 454 AA.  
 AC 055278;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE MLC36 precursor (Fragment).  
 OS Streptococcus sp. "Group C".  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=33972;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C36;  
 RA Ben Nasr A., Wistedt A., Ringdahl U., Sjobring U.;  
 RT "Streptokinase activates plasminogen bound to human group C and group  
 RT G streptococci through M-like proteins.";  
 RL Eur. J. Biochem. 0:0-0(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C36;  
 RX MEDLINE=94291620; PubMed=8020466;  
 RA Ben Nasr A., Wistedt A., Ringdahl U., Sjobring U.;  
 RT "Streptokinase activates plasminogen bound to human group C and group  
 RT G streptococci through M-like proteins.";  
 RL Eur. J. Biochem. 222:267-276(1994).  
 DR EMBL; Z32677; CAA83588.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF02370; M\_4.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 KW Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 29 Potential.  
 FT CHAIN 30 >454 MLC36.  
 FT NON\_TER 454 454  
 SQ SEQUENCE 454 AA; 51416 MW; FA7D34562548282P CRC64;

Query Match 100.0%; Score 64; DB 2; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 0.17;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASREAKKQVEKALE 14  
 Db 332 ASREAKKQVEKALE 345

RESULT 33  
 ID 0840T7 PRELIMINARY; PRT; 454 AA.  
 AC 0840T7;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE M protein (Fragment).  
 GN Name=emm;  
 OS Streptococcus pyogenes.  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dyal-l-Smith M.L., Sriprakash K.S.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY263387; AAO92603.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003345; M\_repeat.



DR Pfam; PF02370; M; 1.  
FT NON\_TER 1  
RP SEQUENCE 454 AA; 51368 MW; 4DP100DAA6467864 CRC64;  
SQ  
Query Match 100.0%; Score 64; DB 2; Length 454;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ASREAKKOVERKALE 14  
Db 402 ASREAKKOVERKALE 415  
RESULT 34  
ID 054510 PRELIMINARY; PRT; 457 AA.  
AC 054510;  
DT 01-NOV-1996 (TRENBLREL. 01, Created)  
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
DE M5\_8193 protein (Fragment).  
GN Name=emm5.8193;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OX Streptococcus.  
NC NCB1\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC8193;  
RX MEDLINE=94224154; PubMed=8170398;  
RA Whatmore A.M., Kehoe M.A.;  
RT "Horizontal gene transfer in the evolution of group A streptococcal  
RL M01. Microbiol. 11:363-374(1994).  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
an amide bond (By similarity).  
DR EMBL; U02480; AA50854.1; -;  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR PRINTS; PR00015; GPOSANCHOR.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor.  
FT NON\_TER 1  
SQ SEQUENCE 457 AA; 51172 MW; DEBBAF50AB6379 CRC64;  
Query Match 100.0%; Score 64; DB 2; Length 457;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ASREAKKOVERKALE 14  
Db 308 ASREAKKOVERKALE 321  
RESULT 35  
ID 083XW0 PRELIMINARY; PRT; 465 AA.  
AC 083XW0;  
DT 01-JUN-2003 (TRENBLREL. 24, Created)  
DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)  
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
CC Streptococcus.

OX NCB1\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22894661; PubMed=14532252;  
RA Lau S.K., Woo P.C., Yim T.C., To A.P., Yuen K.Y.;  
RT "Molecular characterization of a strain of group A streptococcus  
RT isolated from a patient with a posas abscess."  
RL J. Clin. Microbiol. 41:4888-4891(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lau S.K.P., Woo P.C.Y., Yim T.C., To A.P.C., Yuen K.Y.;  
RL Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
an amide bond (By similarity).  
DR EMBL; AY273148; AAP32473.1; -;  
DR HSSP; P04268; IIC2.  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR011000; Apolp\_III\_like.  
DR InterPro; IPR005877; Gpos\_Ysirk.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR PRINTS; PR00015; GPOSANCHOR.  
DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor.  
FT NON\_TER 1  
SQ SEQUENCE 465 AA; 52098 MW; 59717975A94C6B78 CRC64;  
Query Match 100.0%; Score 64; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ASREAKKOVERKALE 14  
Db 327 ASREAKKOVERKALE 340  
RESULT 36  
ID 093SL9 PRELIMINARY; PRT; 471 AA.  
AC 093SL9;  
DT 01-DEC-2001 (TRENBLREL. 19, Created)  
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
DE Emm type protein precursor.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
CC Streptococcus.  
OX NCB1\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rantty R., Jamal F., Ali A.M., Yusoff K.;  
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
an amide bond (By similarity).  
DR EMBL; AY033333; AAK50339.1; -;  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR005877; Gpos\_Ysirk.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF02370; M; 2-  
DR PRINTS; PR00015; Ysirk\_signal; 1.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.

```

DR PROSITE; PS50847; GRAM_POS ANCHORING; 1.
KW Cell wall, peptidoglycan-anchor; Signal.
FT SIGNAL 1 42 Potential.
FT CHAIN 43 471 emm type protein.
SQ SEQUENCE 471 AA; 52636 MW; 72CAEB19FBIADDF9 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREKKOVERKALE 14
Db 322 ASREKKOVERKALE 335

RESULT 37
Q55279 PRELIMINARY; PRT; 472 AA.
AC Q55279;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M1G72 precursor.
OS Streptococcus sp. (lancefield group G).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G72;
RX MEDLINE=94291620; PubMed=8020466;
RA Ben Naer A., Wisniedt A., Ringdahl U., Sjobring U.;
RT "Streptokinase activates plasminogen bound to human group C and group
RT G streptococci through M-like proteins.";
RL Eur. J. Biochem. 222:267-276(1994).
DR EMBL; Z32678; CAA83589.1; -.
DR HSSP; Q9UNH0; 1G04.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M_4.
DR Pfam; PF04650; YsIRK_signal; 1.
KW Signal.
FT SIGNAL 1 29 Potential.
FT CHAIN 30 472 Mature M1G72.
SQ SEQUENCE 472 AA; 53968 MW; 8DE0886B2F45FFC8 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREKKOVERKALE 14
Db 350 ASREKKOVERKALE 363

RESULT 38
Q33631 PRELIMINARY; PRT; 475 AA.
AC Q33631;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE M-like protein.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=25287, and T-type 4;
RC MEDLINE=20123454; PubMed=10660058;
RA Geyer A., Schmidt K.H.;

```

```

RT "Genetic organisation of the M protein region in human isolates of
RT group C and G streptococci: two types of multigene regulator-like
RT (mgc) regions.";
RL Mol. Gen. Genet. 262:965-976(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=25287, and T-type 4;
RX MEDLINE=99447564; PubMed=10518039; DOI=10.1016/S0928-8244(99)00115-7;
RA Geyer A., Roch A., Vettermann S., Guenther E., Groh A., Straube E.,
RA Schmidt K.H.;
RT "M protein of a Streptococcus dysgalactiae human wound isolate shows
RT multiple binding to different plasma proteins and shares epitopes with
RT keratin and human cartilage.";
RL FEWS Immunol. Med. Microbiol. 26:11-24(1999).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; X93464; CAA63750.1; -.
DR EMBL; Y18363; CAB70608.1; -.
DR HSSP; O15813; 1D7M.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_2.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRfams; TIGR01167; LPTX_anchor; 1.
DR TIGRfams; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS ANCHORING; 1.
KW Cell wall, peptidoglycan-anchor.
SQ SEQUENCE 475 AA; 52399 MW; DC077E86DA51EC98 CRC64;

Query Match 100.0%; Score 64; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASREKKOVERKALE 14
Db 320 ASREKKOVERKALE 333

RESULT 39
M6_STRPY STANDARD; PRT; 483 AA.
AC P06089;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE M protein, serotype 6 precursor.
GN Name=emm6;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8611835; PubMed=3511046;
RA Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT "Complete nucleotide sequence of type 6 M protein of the group A
RT Streptococcus. Repetitive structure and membrane anchor.";
RL J. Biol. Chem. 261:1677-1686(1986).
RN [2]
RP SEQUENCE OF 43-122 FROM N.A.
RX MEDLINE=85166224; PubMed=3885219;
RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT "Relationship of M protein genes in group A streptococci.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
CC -1- FUNCTION: This protein is one of the different antigenic serotypes
CC of protein M. Protein M is closely associated with virulence of
CC the bacterium and can render the organism resistant to

```

```

CC phagocytosis.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- SIMILARITY: Belongs to the M protein family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M1138; AAA26920.1; -.
DR PIR; A26297; A26297.
DR InterPro; IPR011000; Apolp III like.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003611; Inton_nuc_2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_2.
DR Pfam; PF04650; Ysirk signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR SMART; SM00456; IENR2; 4.
DR TIGRFAMs; TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Antigen; Cell wall; Coiled coil; Peptidoglycan-anchor; Phagocytosis;
KW Repeat; Signal; Virulence.
FT SIGNAL 1 42
FT CHAIN 1 452
FT PROPEP 453 483
FT DOMAIN 69 138
FT REPEAT 69 75
FT REPEAT 76 82
FT REPEAT 83 89
FT REPEAT 90 96
FT REPEAT 97 103
FT REPEAT 104 110
FT REPEAT 111 117
FT REPEAT 118 124
FT REPEAT 125 131
FT REPEAT 132 138
FT REPEAT 139 269
FT DOMAIN 157 269
FT REPEAT 157 181
FT REPEAT 182 206
FT REPEAT 207 231
FT REPEAT 232 256
FT REPEAT 257 281
FT DOMAIN 279 347
FT DOMAIN 348 411
FT DOMAIN 412 448
FT SITE 449 453
FT MOD_RES 452 452
FT SEQUENCE 483 AA; 53472 MW; 68F87F28DB53A448 CRC64;

Query Match 100.0%; Score 64; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 40  
005464

```

ID 005464 PRELIMINARY; PRT; 484 AA.
AC 005464;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE M1.1 protein precursor.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1114;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS190;
RX MEDLINE=93360826; PubMed=8355619;
RA Harbaugh M.P., Pedbielaki A., Huegl S., Cleary P.P.;
RT "Nucleotide substitutions and small-scale insertion produce size and
RT antigenic variation in group A streptococcal M1 protein."
RL Mol. Microbiol. 8:981-991(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
CC EMBL; Z21845; CAI79893.1; -.
CC PIR; S34978; S34978.
CC PIR; S46489; S46489.
CC HSSP; P04268; IIC2.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_3.
DR Pfam; PF04650; Ysirk signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 41
FT CHAIN 42 484
FT SEQUENCE 484 AA; 54341 MW; 2B310C71F8100CF CRC64;

Query Match 100.0%; Score 64; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Search completed: June 13, 2005, 20:57:31  
Job time : 76.2326 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using SW model

Run on: June 13, 2005, 20:05:19 ; Search time 158.488 Seconds  
(without alignments)  
70.769 Million cell updates/sec

Title: US-10-706-275-2

Perfect score: 137

Sequence: 1 KQAEKVKVSRKAKQVEKALEQLEDKVK 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_GeneSeq\_16Dec04:\*

1: GeneSeqp1980s:\*

2: GeneSeqp1990s:\*

3: GeneSeqp2000s:\*

4: GeneSeqp2001s:\*

5: GeneSeqp2002s:\*

6: GeneSeqp2003s:\*

7: GeneSeqp2003bs:\*

8: GeneSeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137	100.0	29	ADK00565	Immunogen
2	137	100.0	45	ADK00571	Immunogen
3	137	100.0	46	ADK00569	Immunogen
4	137	100.0	47	ADK00572	Immunogen
5	137	100.0	47	ADK00570	Immunogen
6	125	91.2	28	AAW04353	Chimaeric
7	123	89.8	28	AAW04354	Chimaeric
8	116	84.7	28	AAW04355	Chimaeric
9	110	80.3	28	AAW04352	Chimaeric
10	96	70.1	28	AAW04351	Chimaeric
11	81	59.1	28	AAW04350	Chimaeric
12	78	56.9	28	AAW04349	Chimaeric
13	72	52.6	234	AAW04348	Chimaeric
14	72	52.6	234	AAW04347	Chimaeric
15	72	52.6	441	AAW04346	Chimaeric
16	72	52.6	441	AAW04345	Chimaeric
17	72	52.6	441	AAW04344	Chimaeric
18	72	52.6	483	AAW04343	Chimaeric
19	72	52.6	484	AAW04342	Chimaeric
20	72	52.6	484	AAW04341	Chimaeric
21	71	51.8	28	AAW04340	Chimaeric
22	71	51.8	236	AAW04339	Chimaeric
23	71	51.8	236	AAW04338	Chimaeric
24	71	51.8	254	AAW04337	Chimaeric
25	71	51.8	284	AAW04336	Chimaeric
			305	AAW04335	Chimaeric

26	71	51.8	305	3	AAW04334	Chimaeric
27	71	51.8	539	8	AAW04333	Chimaeric
28	68.5	50.0	28	2	AAW04332	Chimaeric
29	67	48.9	29	2	AAW04331	Chimaeric
30	67	48.9	389	6	AAW04330	Chimaeric
31	66	48.2	284	8	AAW04329	Chimaeric
32	65.5	47.8	28	2	AAW04328	Chimaeric
33	65	47.4	29	2	AAW04327	Chimaeric
34	64	46.7	20	2	AAW04326	Chimaeric
35	64	46.7	20	2	AAW04325	Chimaeric
36	64	46.7	20	2	AAW04324	Chimaeric
37	64	46.7	29	2	AAW04323	Chimaeric
38	64	46.7	37	6	AAW04322	Chimaeric
39	64	46.7	72	2	AAW04321	Chimaeric
40	64	46.7	361	7	AAW04320	Chimaeric
41	63.5	46.4	65	7	AAW04319	Chimaeric
42	63	46.0	29	2	AAW04318	Chimaeric
43	63	46.0	29	2	AAW04317	Chimaeric
44	63	46.0	107	2	AAW04316	Chimaeric
45	63	46.0	107	2	AAW04315	Chimaeric

## ALIGNMENTS

RESULT 1	ADK00565	standard; peptide; 29 AA.
XX	ADK00565	
AC	ADK00565	
DT	06-MAY-2004	(first entry)
DE	Immunogenic lipopeptide of the invention #101.	
DE	T helper cell epitope; B cell epitope; Antibacterial; Anticancer;	
KW	Antiferility; Vaccine; antibody.	
KW	Synthetic.	
OS	WO2004014956-A1.	
PN	19-FEB-2004.	
PD	12-AUG-2003; 2003WO-AU001018.	
PF	12-AUG-2002; 2002US-0402838P.	
PR	(COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.	
PA	Jackson D, Zeng W;	
XX	WPI; 2004-238735/22.	
DR	Novel lipopeptide comprising polypeptide having amino acid sequence of T	
PT	helper cell epitope and B cell epitope, conjugated to lipid moieties,	
PT	useful for eliciting immune response against group A Streptococcus	
PT	antigen.	
XX	Claim 29; SEQ ID NO 101; 194bp; English.	
PS	The present invention relates to a lipopeptide comprising polypeptide	
CC	conjugated to lipid moieties, where polypeptide contains amino acid	
CC	sequence of T helper cell epitope and B cell epitope, where amino acid	
CC	sequences are different, and internal lysine residues or internal lysine	
CC	analog residues for covalent attachment of each of lipid moieties through	
CC	key; amino group or terminal side chain group of lysine or lysine	
CC	analog. The peptides are useful in eliciting the production of antibody	
CC	against an antigenic B cell epitope in a subject, and are useful for	
CC	antibody production, synthetic vaccine production, diagnostic method	
CC	employing antibodies and antibody ligands and immunotherapy for	
CC	veterinary and human medicine. The method efficiently elicits the	
CC	production of antibody against antigenic B cell epitope. The present	



DE Immunogenic lipopeptide of the invention #108.  
XX  
XX T helper cell epitope; B cell epitope; Antibacterial; Antitumor;  
KW Antifertility; Vaccine; antibody.  
XX  
OS Synthetic.  
XX WO2004014956-A1.  
XX  
XX 19-FEB-2004.  
XX  
XX 12-AUG-2003; 2003WO-AU001018.  
XX  
XX 12-AUG-2002; 2002US-0402838P.  
XX  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
XX Jackson D, Zeng W;  
XX  
XX WPI; 2004-238735/22.  
XX  
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.  
XX  
XX Claim 39; SEQ ID NO 108; 194pp; English.  
XX  
XX The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC ketyl; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.  
XX  
XX  
SQ Sequence 46 AA;  
Query Match 100.0%; Score 137; DB 8; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KOAEDKVASREAKKQVEKALEQLEDKYK 29  
Db 18 KOAEDKVASREAKKQVEKALEQLEDKYK 46  
RESULT 5  
ADK00570  
ID ADK00570 standard; peptide; 47 AA.  
XX  
XX ADK00570;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX Immunogenic lipopeptide of the invention #106.  
XX  
XX T helper cell epitope; B cell epitope; Antibacterial; Antitumor;  
KW Antifertility; Vaccine; antibody.  
XX  
XX Synthetic.  
XX WO2004014956-A1.  
XX  
XX 19-FEB-2004.  
XX  
XX

PF 12-AUG-2003; 2003WO-AU001018.  
XX  
XX  
XX 12-AUG-2002; 2002US-0402838P.  
XX  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
XX Jackson D, Zeng W;  
XX  
XX WPI; 2004-238735/22.  
XX  
XX  
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.  
XX  
XX Claim 39; SEQ ID NO 106; 194pp; English.  
XX  
XX The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC ketyl; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.  
XX  
XX  
SQ Sequence 47 AA;  
Query Match 100.0%; Score 137; DB 8; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KOAEDKVASREAKKQVEKALEQLEDKYK 29  
Db 19 KOAEDKVASREAKKQVEKALEQLEDKYK 47  
RESULT 6  
AAW04353  
ID AAW04353 standard; procein; 28 AA.  
XX  
XX AAW04353;  
XX  
XX 02-DEC-1996 (first entry)  
XX  
XX Chimaeric peptide (U7) contg. Streptococcal M protein peptide p145.  
XX  
XX Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCM4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
XX diagnosis.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH 15..26  
FT /note="p145 conformational B-cell epitope"  
XX  
XX WO9611944-A1.  
XX  
XX 25-APR-1996.  
XX  
XX 16-OCT-1995; 95WO-AU000681.  
XX  
XX 14-OCT-1994; 94AU-00008851.  
XX  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX  
XX

PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX  
DR WPI; 1996-221939/22.  
XX  
PT New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 12; Fig 1C; 99pp; English.  
XX  
CC The present peptide is a chimaeric peptide (CP) contg. the Streptococcal  
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
CC (1992)). The CP comprises a B-cell conformational epitope from within  
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
CC the GCN4 leucine zipper peptide (AA97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
XX  
SQ Sequence 28 AA;  
XX  
Query Match 91.2%; Score 125; DB 2; Length 28;  
Best Local Similarity 92.9%; Pred. No. 3e-08;  
Matches 26; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QAEEDKVKASREAKKQVEKALEQLEDDKY 28  
Db 1 QAEEDKVKASREAKKQVEKAVKQLEDDKY 28  
XX  
RESULT 7  
AAW04354  
ID AAW04354 standard; protein; 28 AA.  
XX  
AC AAW04354;  
XX  
DT 02-DEC-1996 (first entry)  
XX  
DE Chimaeric peptide (U8) contg. Streptococcal M protein peptide p145.  
XX  
KW Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KW diagnosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 16..27  
/note="p145 conformational B-cell epitope"  
XX  
PN W09611944-A1.  
XX  
PD 25-APR-1996.  
XX  
PE 16-OCT-1995; 95WO-AU000681.  
XX  
PR 14-OCT-1994; 94AU-00008851.  
XX  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX  
DR WPI; 1996-221939/22.  
XX  
PT New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 12; Fig 1C; 99pp; English.  
XX  
CC The present peptide is a chimaeric peptide (CP) contg. the Streptococcal  
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
CC (1992)). The CP comprises a B-cell conformational epitope from within  
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
CC the GCN4 leucine zipper peptide (AA97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
XX  
SQ Sequence 28 AA;  
XX  
Query Match 89.8%; Score 123; DB 2; Length 28;  
Best Local Similarity 92.9%; Pred. No. 5.3e-08;  
Matches 26; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 QAEEDKVKASREAKKQVEKALEQLEDDKY 29  
Db 1 QAEEDKVKASREAKKQVEKALKQLEDDKY 28  
XX  
RESULT 8  
AAW04355  
ID AAW04355 standard; protein; 28 AA.  
XX  
AC AAW04355;  
XX  
DT 02-DEC-1996 (first entry)  
XX  
DE Chimaeric peptide (U9) contg. Streptococcal M protein peptide p145.  
XX  
KW Streptococcal; M protein; peptide; p145; chimaeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KW diagnosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 17..28  
/note="p145 conformational B-cell epitope"  
XX  
PN W09611944-A1.  
XX  
PD 25-APR-1996.  
XX  
PE 16-OCT-1995; 95WO-AU000681.  
XX  
PR 14-OCT-1994; 94AU-00008851.  
XX  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.



PA (CSLC-) CSL LTD.  
XX  
PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX WPI; 1996-221939/22.  
DR  
XX New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 12; Fig 1C; 99pp; English.  
XX  
CC The present peptide is a chimeric peptide (CP) contg. the Streptococcal  
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
CC (1992)). The CP comprises a B-cell conformational epitope from within  
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
XX  
SQ Sequence 28 AA;  
XX  
Query Match 84.7%; Score 116; DB 2; Length 28;  
Best Local Similarity 92.6%; Pred. No. 3.7e-07;  
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
QY 3 AEDKVKASREAKKQVEKALEQLEDEKVK 29  
Db 1 AEDKVKQUREAKKQVEKALEQLEDEKVK 27  
XX  
RESULT 9  
AAW04352  
ID AAW04352 standard; protein; 28 AA.  
XX  
AC AAW04352;  
XX  
DT 02-DEC-1996 (first entry)  
XX  
DE Chimeric peptide (J6) contg. Streptococcal M protein peptide p145.  
XX  
KW Streptococcal; M protein; peptide; p145; chimeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KW diagnosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 14..25  
FT /note="p145 conformational B-cell epitope"  
XX  
XX W09611944-A1.  
XX  
XX 25-APR-1996.  
XX  
XX 16-OCT-1995; 95WO-AU000681.  
XX  
XX 14-OCT-1994; 94AU-00008851.  
XX  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.  
XX (UYME) UNIV MELBOURNE.  
XX (HALL-) HALT. INST MEDICAL RES WALTER & ELIZA.  
XX (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
XX (CSLC-) CSL LTD.  
XX

PI Cooper JA, Relf WA, Good MF, Saul AJ;  
XX WPI; 1996-221939/22.  
DR  
XX New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
XX  
PS Example 12; Fig 1C; 99pp; English.  
XX  
CC The present peptide is a chimeric peptide (CP) contg. the Streptococcal  
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735  
CC (1992)). The CP comprises a B-cell conformational epitope from within  
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on  
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar  
CC conformation, enabling the epitope to be presented in an immunologically  
CC active conformation. The CP can be used in a novel detection/mapping  
CC process, e.g. to determine the min. epitope required to induce opsonic  
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised  
CC against the CP can be used for immunotherapy and diagnosis, while the CP  
CC can be used diagnostically to detect Ab. The reactivity of the present  
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance  
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145  
XX  
SQ Sequence 28 AA;  
XX  
Query Match 80.3%; Score 110; DB 2; Length 28;  
Best Local Similarity 85.2%; Pred. No. 2e-06;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
XX  
QY 1 KOAEDKVKASREAKKQVEKALEQLEDEK 27  
Db 2 KOAEDKVDASREAKKQVEKVKQLEDEK 28  
XX  
RESULT 10  
AAW04351  
ID AAW04351 standard; protein; 28 AA.  
XX  
AC AAW04351;  
XX  
DT 02-DEC-1996 (first entry)  
XX  
DE Chimeric peptide (J5) contg. Streptococcal M protein peptide p145.  
XX  
KW Streptococcal; M protein; peptide; p145; chimeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KW diagnosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 13..24  
FT /note="p145 conformational B-cell epitope"  
XX  
XX W09611944-A1.  
XX  
XX 25-APR-1996.  
XX  
XX 16-OCT-1995; 95WO-AU000681.  
XX  
XX 14-OCT-1994; 94AU-00008851.  
XX  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.  
XX (UYME) UNIV MELBOURNE.  
XX (HALL-) HALT. INST MEDICAL RES WALTER & ELIZA.  
XX (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
XX (CSLC-) CSL LTD.  
XX  
XX Cooper JA, Relf WA, Good MF, Saul AJ;  
XX

```

DR WPI, 1996-221939/22.
XX
XX New chimeric peptide(s) including a conformational epitope - inserted
PT into a peptide having similar native conformation, useful in vaccines and
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.
XX
XX Example 12; Fig 1C; 99pp; English.
XX
CC The present peptide is a chimaeric peptide (CP) contg. the Streptococcal
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735
CC (1992)). The CP comprises a B-cell conformational epitope from within
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar
CC conformation, enabling the epitope to be presented in an immunologically
CC active conformation. The CP can be used in a novel detection/mapping
CC process, e.g. to determine the min. epitope required to induce opsonic
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised
CC against the CP can be used for immunotherapy and diagnosis, while the CP
CC can be used diagnostically to detect Ab. The reactivity of the present
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145
XX
XX Sequence 28 AA;
SQ
Query Match 70.1%; Score 96; DB 2; Length 28;
Best Local Similarity 76.9%; Pred. No. 9.9e-05;
Matches 20; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Oy 1 KOAEDKVKASREAKKQVEKALEQLE 26
Db 3 KOAEDKLDASREAKKQVEDKVKQLE 28
RESULT 11
AAM04350 ID AAM04350 standard; protein; 28 AA.
XX
XX AAM04350;
AC
XX 02-DEC-1996 (first entry)
DT
XX Chimaeric peptide (V4) contg. Streptococcal M protein peptide p145.
DE
XX Streptococcal M protein; peptide; p145; chimaeric; chimeric; B-cell;
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;
KM mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;
XX diagnosis.
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FT Peptide 12..23
FT /note="p145 conformational B-cell epitope"
XX
XX WO9611944-A1.
XX
XX 25-APR-1996.
XX
XX 16-OCT-1995; 95WO-AU000681.
XX
XX 14-OCT-1994; 94AU-00008851.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX (UYME) UNIV MELBOURNE.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX (BIOT-) BIOTECH AUSTRALIA PTY LTD.
XX (CSLC-) CSL LTD.
XX
XX Cooper JA, Relf WA, Good MF, Saul AJ;
XX
XX WPI, 1996-221939/22.
XX

```

```

PT New chimeric peptide(s) including a conformational epitope - inserted
PT into a peptide having similar native conformation, useful in vaccines and
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.
XX
XX Example 12; Fig 1C; 99pp; English.
XX
CC The present peptide is a chimaeric peptide (CP) contg. the Streptococcal
CC M protein peptide p145 (Pruksakorn et al, J. Immunol. 149: 2729-2735
CC (1992)). The CP comprises a B-cell conformational epitope from within
CC p145, inserted into a 2nd peptide, pref. an alpha-helical coil based on
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a similar
CC conformation, enabling the epitope to be presented in an immunologically
CC active conformation. The CP can be used in a novel detection/mapping
CC process, e.g. to determine the min. epitope required to induce opsonic
CC antibodies (Ab), and in vaccines against gp. A Streptococci. Ab raised
CC against the CP can be used for immunotherapy and diagnosis, while the CP
CC can be used diagnostically to detect Ab. The reactivity of the present
CC peptide to the anti-p145 mouse serum 11/1(8) given as a mean absorbance
CC value at 405 nm was ca. 0.2, compared to ca. 1.8 for p145
XX
XX Sequence 28 AA;
SQ
Query Match 59.1%; Score 81; DB 2; Length 28;
Best Local Similarity 68.0%; Pred. No. 0.0065;
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Oy 1 KOAEDKVKASREAKKQVEKALEQLE 25
Db 4 KOAEDLDASREAKKQVEDKVKQLE 28
RESULT 12
AAR97454 ID AAR97454 standard; peptide; 28 AA.
XX
XX AAR97454;
AC
XX 04-DEC-1996 (first entry)
DT
XX Chimaeric peptide av86 contg. C. elegans unc-15 paramyosin peptide.
DE
XX Paramyosin protein; peptide; unc-15; chimaeric; chimeric; B-cell;
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;
KM mapping; opsonic antibody; vaccine; immunotherapy; diagnosis.
XX Synthetic.
OS
XX
XX WO9611944-A1.
XX
XX 25-APR-1996.
XX
XX 16-OCT-1995; 95WO-AU000681.
XX
XX 14-OCT-1994; 94AU-00008851.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX (UYME) UNIV MELBOURNE.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX (BIOT-) BIOTECH AUSTRALIA PTY LTD.
XX (CSLC-) CSL LTD.
XX
XX Cooper JA, Relf WA, Good MF, Saul AJ;
XX
XX WPI, 1996-221939/22.
XX
XX New chimeric peptide(s) including a conformational epitope - inserted
XX into a peptide having similar native conformation, useful in vaccines and
XX for determ. of minimal epitope(s) or for mapping amphipathic helices.
XX
XX Example 18; Page 44; 99pp; English.
XX
XX The present peptide is a chimaeric peptide (CP), contg. a C. elegans unc-
CC

```

CC 15 paramyosin peptide. The CP comprises a B-cell conformational epitope  
 CC from within unc-15, inserted into a 2nd peptide, pref. an alpha-helical  
 CC based on the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide  
 CC has a similar conformation, enabling the epitope to be presented in an  
 CC immunologically active conformation. The CP can be used in a novel  
 CC detection/mapping process, e.g. to determine the min. epitope required to  
 CC induce opsonic antibodies (Ab), and in vaccines against C. elegans. Ab  
 CC raised against the CP can be used for immunotherapy and diagnosis, while  
 CC the CP can be used diagnostically to detect Ab

SO Sequence 28 AA;

Query Match 56.9%; Score 78; DB 2; Length 28;  
 Best Local Similarity 64.0%; Pred. No. 0.015;  
 Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KOAEKVKASREAKKQVEKALEOLE 25  
 DB 4 KOAEDDLASREAKKQLODKVKOLE 28

RESULT 13

ID AAR10221 standard; protein; 234 AA.

XX AAR10221;

XX 25-MAR-2003 (revised)

DT 26-MAR-1991 (first entry)

XX Streptococcal M' protein.

XX Streptococcal M protein; M' 6 protein; vaccinia virus; fowlpox virus;

KM poxviridae vaccine; streptococcal pharyngitis.

XX Streptococcus sp.

XX WO9015872-A.

XX 27-DEC-1990.

XX 21-JUN-1989; 89US-00369118.

XX 21-JUN-1989; 89US-00369118.

PR 19-JUN-1990; 90US-00540586.

XX (UYRO ) UNIV ROCKEFELLER.

PA (UYOR-) UNIV OREGON STATE.

XX Fischetti VA, Kraby DE;

XX WPI, 1991-022236/03.

DR N-PSDB; AAQ10244.

PT New recombinant streptococcal M protein DNA and viral vector - for  
 PT production of poxviridae vaccines in treatment of vaccinia, fowlpox etc.

XX Disclosure; Fig 5; 41pp; English.

XX This M' 6 protein corresponds to the conserved exposed polypeptide region  
 CC of the streptococcal M protein. It is encoded by a gene- tically  
 CC engineered gene introduced into the genome of a vaccinia or fowlpox  
 CC virus. The resultant DNA complex is useful as a vaccine for  
 CC immunoprotection against streptococcal infections. The M' 6 polypeptide is  
 CC the part of protein M responsible for virulence. (Updated on 25-MAR-2003  
 CC to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

SO Sequence 234 AA;

Query Match 52.6%; Score 72; DB 2; Length 234;  
 Best Local Similarity 45.2%; Pred. No. 0.81;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKKQVEKALEOLEDKV 28  
 DB 63 DKVKEKQISDASRQGLRRDDASREAKKQVEKALEANSKTL 104

RESULT 14

ID AAR20128 standard; protein; 281 AA.

XX AAR20128;

XX 27-AUG-2003 (revised)

DT 15-APR-1992 (first entry)

XX Sequence encoded by truncated M1 gene.

XX Protein H; immunoglobulin G; IgG; antibody; autoimmune disease.

XX Streptococcus sp.

XX Key Location/Qualifiers

FT Region 1-28

FT /label= p16M1

FT Region 29..70

FT /label= C1

FT Region 71..112

FT /label= C2

FT Region 113..155

FT /label= C3

FT Region 156..176

FT /label= C4

FT Region 177..281

FT /label= D

XX WO9119740-A.

XX 26-DEC-1991.

XX 21-JUN-1990; 90SE-00002212.

XX 21-JUN-1990; 90SE-00002212.

XX (HIGH-) HIGHTECH RECEPTOR A.

XX Schmidt KH, Akesson P, Cooney J, Bjorck L;

XX WPI, 1992-024366/03.

DR N-PSDB; AAQ20292.

PT New IgG binding proteins H' lacking an albumin binding sequence - useful  
 PT in purificn. of excess IgG from blood and to diagnose autoimmune  
 PT diseases.

XX Disclosure; Fig 8; 37pp; English.

XX The inventors claim a protein prod. by a strain of Gp. A Streptococci. The  
 CC protein has the AA sequence of protein H but lacks at least some part of  
 CC the C and D regions (responsible for binding albumin), esp. it lacks the  
 CC whole of these regions and extends for AA1 to AA158. Compared with  
 CC natural protein H, it is more specific and may be used as part of a kit  
 CC for the binding, separation and identification of human IgG. The same  
 CC sequences appear in WO9119741. (Updated on 27-AUG-2003 to correct OS  
 CC field.)

SO Sequence 281 AA;

Query Match 52.6%; Score 72; DB 2; Length 281;  
 Best Local Similarity 45.2%; Pred. No. 0.99;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
 QY 5 DKVK-----ASREAKKQVEKALEOLEDKV 28  
 DB 110 DKVKEKQISDASRQGLRRDDASREAKKQVEKALEANSKTL 151



```

XX DE Streptococcus sp. M6 protein.
XX KM Antigen conjugate; immune response; infection; C repeat region;
XX KM cholera toxin B; CTB.
XX OS Streptococcus sp.
XX PN US6602507-B1.
XX PD 05-AUG-2003.
XX PF 06-JAN-1995; 95US-00369295.
XX PR 25-MAR-1988; 88US-00173380.
XX PR 27-FEB-1988; 89US-00315588.
XX PR 19-JUN-1990; 90US-00340101.
XX PR 03-MAR-1992; 92US-00845865.
XX PR 28-MAY-1993; 93US-00068598.
XX PR 22-FEB-1994; 94US-00200913.
XX PA (UNIQ ) UNIV ROCKEFELLER.
XX P1 Fischetti VA;
XX PI MPI; 2003-895270/82.
XX PT New antigen conjugates comprising polypeptides from the conserved exposed
XX PR region of group A streptococcal M protein, useful for preparing vaccines
XX PT for controlling streptococcal infection a mammal.
XX PS Disclosure; Fig 1; 12pp; English.
XX CC The invention relates an antigen conjugate comprising a linkable carrier
XX CC covalently bound to a polypeptide consisting of five or more amino acid
XX CC residues from the conserved exposed region of group A streptococcal M
XX CC protein. The antigen conjugate elicits a protective immune response to
XX CC streptococcal infection in a mammal when administered mucosally. The
XX CC antigen conjugates and polypeptides are useful for preparing vaccines for
XX CC controlling streptococcal infection or protecting against streptococcal
XX CC infection in a mammal. The polypeptides were all within the C repeat
XX CC region of the M protein of type 6 Streptococci (strain 471). The
XX CC polypeptides were purified by high performance liquid chromatography.
XX CC Highly purified cholera toxin B (CTB) was derivatised at primary amino
XX CC groups with the heterobifunctional cross-linking agent N-succinimidyl 3-
XX CC (2-pyridyldithio) propionate (SPDP). The synthetic peptides were mixed
XX CC with SPDP-derivatised CTB at a 1.5:1 ratio. An average of 1.06-1.27
XX CC peptide molecules were covalently linked per CTB monomer. The present
XX CC sequence is M6 protein from strain DA71 of group A Streptococcus sp.
XX SQ Sequence 441 AA;
SQ Query Match 52.6%; Score 72; DB 7; Length 441;
Best Local Similarity 45.2%; Pred. No. 1.6;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;
QY 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 270 DKVKEKQISDSRQGLRRDLDSREAKKQVEKALEBANSKL 311
RESULT 18
AAW08927
ID AAW08927 standard; protein; 483 AA.
XX AC AAW08927;
XX DT 31-MAR-1997 (first entry)
XX DE Type-6 M-protein.
XX KM Type-6 M-protein; emm6 gene; Streptococcus pyogenes; signal peptide;
KM anchor peptide; C-terminal sorting signal; surface display;

```

```

KM fusion protein; protease-deficient; Streptococcus gordonii; SPEX vector;
KM fusion protein cleavage; spacer; protein secretion; TBY-N1A protease;
XX KM diagnostic; recombinant vaccine; therapy.
XX OS Streptococcus pyogenes.
XX PN Key
XX FH Location/Qualifiers
XX FT 1..58
XX FT /note= "Preferred N-terminal secretion sequence (claim
XX FT 23)"
XX FT Peptide
XX FT 1..42
XX FT /note= "Signal peptide (claim 22)"
XX FT Peptide
XX FT 263..342
XX FT /note= "Preferred spacer peptide (claim 32)"
XX FT Peptide
XX FT 344..483
XX FT /note= "Preferred C-terminal anchoring sequence (claim
XX FT 25)"
XX FT Domain
XX FT 405..458
XX FT /note= "Proline/glycine-rich cell wall domain"
XX FT Peptide
XX FT 449..454
XX FT /note= "Conserved C-terminal motif (AAW08924, claim 9)"
XX FT Domain
XX FT 459..477
XX FT /note= "Hydrophobic membrane-spanning domain"
XX FT 478..483
XX FT /note= "C-terminal charged tail"
XX PN WO9640943-A1.
XX PD 19-DEC-1996.
XX PF 06-JUN-1996; 96WO-US009965.
XX PR 07-JUN-1995; 95US-00472244.
XX PA (UNIQ ) UNIV ROCKEFELLER.
XX PI Darzins A, Whitehead S, Hruby D;
XX PT MPI; 1997-052336/05.
XX PT N-PSDB; MAT49317.
XX PR Expressing heterologous proteins in gram-positive bacteria - produces
XX PR proteins anchored to host's cell surface, or secreted from cell.
XX PS Disclosure; Fig 3; 41pp; English.
XX CC This sequence represents Streptococcus pyogenes type-6 M-protein, which
XX CC is a cell wall surface protein. The N-terminal signal peptide, part of
XX CC the mature N-terminal sequence, the C-terminal sorting signal (anchor)
XX CC sequence and optionally a spacer region from the protein may be expressed
XX CC as a fusion with a target protein in a protease-deficient Gram-positive
XX CC host, e.g. the human oral commensal Streptococcus gordonii, in a new
XX CC cloning method involving the SPEX vector series (e.g. plasmid pSPEX1a).
XX CC The C-terminal sorting signal contains a proline/glycine-rich region
XX CC (with a conserved cell wall anchoring motif, AAW08924), a hydrophobic
XX CC membrane-spanning region and a charged tail. The resulting fusion protein
XX CC is expressed on the cell surface, and may be cleaved with e.g. tobacco-
XX CC etch virus TBY-N1A protease (at a site adjacent to the anchor, e.g.
XX CC AAW08925-26) for simple purification. A spacer region from M-protein may
XX CC be included to reduce steric hindrance during fusion protein cleavage. A
XX CC vector without the anchor sequence may be used for protein secretion. The
XX CC method may be used in production of recombinant diagnostic, vaccine and
XX CC therapeutic proteins
XX SQ Sequence 483 AA;
SQ Query Match 52.6%; Score 72; DB 2; Length 483;
Best Local Similarity 45.2%; Pred. No. 1.8;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;
QY 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 312 DKVKEKQISDSRQGLRRDLDSREAKKQVEKALEBANSKL 353

```

## RESULT 19

ABP30015  
ID ABP30015 standard; protein; 484 AA.  
XX  
XX  
AC ABP30015;  
XX  
XX 02-JUL-2002 (first entry)  
DT  
XX  
XX Streptococcus polypeptide SEQ ID NO 9206.  
DE  
XX  
XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX  
XX Streptococcus pyogenes.  
OS  
XX  
XX WO200234771-A2.  
PN  
XX  
XX 02-MAY-2002.  
PD  
XX  
XX 29-OCT-2001; 2001WO-GB004789.  
PF  
XX  
XX 27-OCT-2000; 2000GB-00026333.  
PR 24-NOV-2000; 2000GB-00028727.  
PR 07-MAR-2001; 2001GB-00005640.  
XX  
XX (CHTR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Telford J, Maignan V, Margalit Y Roel, Grandi G, Fraser C;  
PI Tettelin H;  
XX  
XX WPI: 2002-352536/38.  
DR N-PSDB; AEN70646.  
XX  
XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.  
XX  
XX Claim 1; Page 4044; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), AEN6044-AEN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins  
XX  
XX  
SQ Sequence 484 AA;

Query Match 52.6%; Score 72; DB 5; Length 484;  
Best Local Similarity 45.2%; Pred. No. 1.8;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
|||  
|||  
DB 313 DKVKEKQISDASRQGLRDLDAAREAKQVEKALEANSKL 354

RESULT 20

ADR83969  
ID ADR83969 standard; protein; 484 AA.  
XX  
XX  
AC ADR83969;  
XX  
XX 02-DEC-2004 (first entry)  
DT  
XX  
XX S. pyogenes hyperimmune system reactive antigen Spy2018.  
DE  
XX  
XX hyperimmune serum reactive antigen; vaccine; anticaline.  
KW  
XX  
XX Streptococcus pyogenes.  
OS  
XX  
XX WO2004078907-A2.  
PN  
XX  
XX 16-SEP-2004.  
PD  
XX  
XX 02-MAR-2004; 2004WO-EP002087.  
PF  
XX  
XX 04-MAR-2003; 2003EP-00450061.  
PR  
XX  
XX (INTE-) INTERCELL AG.  
PA  
XX  
XX Meinke A, Nagy E, Winkler B, Gelmann D;  
PI  
XX  
XX WPI: 2004-653698/63.  
DR N-PSDB; ADR83819.  
XX  
XX

PT New isolated nucleic acid molecules encoding hyperimmune serum-reactive  
PT antigens from Streptococcus pyogenes, useful for diagnosing, preventing  
PT and treating S. pyogenes infections.  
XX  
XX  
XX Claim 14; SEQ ID NO 237; 145pp; English.  
XX  
XX

CC This invention describes a novel nucleic acid molecule encoding a  
CC hyperimmune serum reactive antigen or its fragment from Streptococcus  
CC pyogenes. The nucleic acid molecule or hyperimmune serum-reactive antigen  
CC or its fragment are useful for the manufacture of a pharmaceutical  
CC preparation, especially a vaccine, against S. pyogenes infection. In  
CC addition, the hyperimmune serum reactive antigen or fragment is used for  
CC the isolation and/or purification and/or identification of an interaction  
CC partner of the hyperimmune serum reactive antigen or its fragment, for  
CC the generation of a peptide (e.g. anticalines) binding to the antigen or  
CC fragment, or for the manufacture of a functional nucleic acid selected  
CC for aptamers and spiegelmers. The nucleic acid molecule may also be used  
CC for the manufacture of functional ribonucleic acids, such as ribozymes.  
CC antisense nucleic acids and siRNA. ADR83733-ADR84189 represent S.  
CC pyogenes hyperimmune serum reactive antigens, fragments and the encoding  
CC polynucleotide described in the invention.  
XX  
XX  
XX  
SQ Sequence 484 AA;

Query Match 52.6%; Score 72; DB 8; Length 484;  
Best Local Similarity 45.2%; Pred. No. 1.8;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
|||  
|||  
DB 313 DKVKEKQISDASRQGLRDLDAAREAKQVEKALEANSKL 354

## RESULT 21

AAR97396  
ID AAR97396 standard; peptide; 28 AA.  
XX  
XX  
AC AAR97396;  
XX  
XX 02-DEC-1996 (first entry)  
DT  
XX  
XX (GCN4) 4 alpha-helical coil used in chimaeric peptide.  
DE  
XX  
XX Streptococcal M protein; peptide; p145; chimeric; B-cell;  
KW conformational epitope; alpha-helix; detection; mapping;

KW opsonic antibody; vaccine; paramyxosin; unc-15; Caenorhabditis elegans;  
 KW group A streptococci; immunotherapy; diagnosis; (GCN4)4.  
 XX Synthetic.  
 OS  
 XX W09611944-A1.  
 XX  
 XX 25-APR-1996.  
 XX  
 XX 16-OCT-1995; 95WO-AU0006681.  
 XX  
 XX 14-OCT-1994; 94AU-00008851.  
 XX  
 XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.  
 XX (UYME) UNIV MELBOURNE.  
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
 XX (CSLC-) CSL LTD.  
 XX  
 XX Cooper JA, Relf WA, Good MF, Saul AJ;  
 XX WPI; 1996-221939/22.  
 XX  
 XX New chimeric peptide(s) including a conformational epitope - inserted  
 XX into a peptide having similar native conformation, useful in vaccines and  
 XX for determ. of minimal epitope(s) or for mapping amphipathic helices.  
 XX  
 XX Example 11; Fig 1A; 99pp; English.  
 XX  
 XX A novel chimeric peptide (CP) comprises a B-cell conformational epitope  
 XX from within the streptococcal M protein peptide p145, or the  
 XX Caenorhabditis elegans paramyxosin unc-15 protein, inserted into a 2nd  
 XX peptide, pref. the alpha-helical coil (GCN4)4. The 2nd peptide has a  
 XX similar conformation to the epitope, enabling the epitope to be presented  
 XX in an immunologically active conformation. The CP can be used in a novel  
 XX detection/mapping process, e.g. to determine the min. epitope required to  
 XX induce opsonic antibodies (Ab), and in vaccines against GP. A  
 XX Streptococci. Ab raised against the CP can be used for immunotherapy and  
 XX diagnosis, while the CP can be used diagnostically to detect Ab  
 XX  
 XX Sequence 28 AA;  
 XX  
 XX Query Match 51.8%; Score 71; DB 2; Length 28;  
 XX Best Local Similarity 55.6%; Pred. No. 0.1;  
 XX Matches 15; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 XX  
 XX QY 1 KOAEDKVASREAKKQVEKALEQLEDX 27  
 XX ||||| : ||| : |||||  
 XX 2 KQLEDKVKQLEDKVKQLEDKVKQLEDX 28  
 XX  
 XX RESULT 22  
 XX AAB03118  
 XX ID AAB03118 standard; protein; 236 AA.  
 XX  
 XX AC AAB03118;  
 XX  
 XX DT 10-OCT-2000 (first entry)  
 XX  
 XX DB C-terminus of Streptococcus pyogenes M protein M5.  
 XX  
 XX KW Multivalent hybrid M protein; M5; group A streptococcus; serotype;  
 KW immunogenic; sero-specific antibody; streptococcal infection;  
 KW cross reactivity; vaccine; acute rheumatic fever; ARF;  
 KW rheumatic heart disease; streptococcal pharyngitis; strep throat;  
 KW pneumonia.  
 XX  
 XX OS Streptococcus pyogenes.  
 XX  
 XX PN US6063386-A.  
 XX  
 XX PD 16-MAY-2000.

XX  
 PF 15-SEP-1997; 97US-00937271.  
 XX  
 XX 16-SEP-1992; 92US-00945954.  
 XX  
 XX (UYTE-) UNIV TENNESSEE RES CORP.  
 XX  
 XX Lederer JW, Dale JB;  
 XX WPI; 2000-364475/31.  
 XX  
 XX New immunogenic recombinant hybrid M protein comprising amino-terminal  
 XX peptide fragments of streptococcal M protein useful as vaccine against  
 XX rheumatic fever and infections leading to rheumatic fever.  
 XX  
 XX Disclosure; Col 45-46; 62pp; English.  
 XX  
 XX The invention relates to multivalent immunogenic hybrid group A  
 XX streptococcal M proteins comprising N-terminal peptide fragments of M  
 XX proteins that elicit opsonic antibodies against multiple serotypes of  
 XX group A streptococci (e.g., Streptococcus pyogenes). The antibodies  
 XX generated using the hybrid proteins are against one or more M protein  
 XX serotypes, where at least one of the serotypes is M1, M3, M18 or M19. The  
 XX invention also encompasses a recombinant DNA molecule comprising a  
 XX nucleotide sequence that encodes a multivalent hybrid M protein; and a  
 XX method for administering a mammal against streptococcal infections,  
 XX comprising administering an immunogenic multivalent hybrid M protein to  
 XX the mammal. The multivalent hybrid M proteins are useful for eliciting  
 XX opsonic or protective antibodies to the M proteins of different serotypes  
 XX of group A streptococci, and may therefore be used as vaccines to protect  
 XX against and control infection by type A streptococci. Type A streptococci  
 XX are not only responsible for streptococcal pharyngitis (strep throat),  
 XX forms of pneumonia and a condition resembling toxic shock, but are also  
 XX involved in the development of acute rheumatic fever (ARF) and rheumatic  
 XX heart disease. In a patient with ARF, antibodies formed during a group A  
 XX streptococcal infection are also cross-reactive with heart tissue, which  
 XX indicates that the streptococci and host tissue contain similar antigenic  
 XX motifs. The new multivalent vaccines are capable of raising sero-specific  
 XX antibodies against various serotypes of group A streptococci which are  
 XX not cross-reactive with human heart tissue. The present sequence  
 XX represents the C-terminal half of the Streptococcus pyogenes M5 protein,  
 XX which is a component of the tetravalent M24-M5-M6-M19 hybrid M protein  
 XX AAB03117  
 XX  
 XX Sequence 236 AA;  
 XX  
 XX Query Match 51.8%; Score 71; DB 3; Length 236;  
 XX Best Local Similarity 75.0%; Pred. No. 1.1;  
 XX Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 XX  
 XX QY 9 ASREAKKQVEKALEQLEDX 28  
 XX ||||| : ||| : |||||  
 XX DB 87 ASREAKKQVEKALEBANSKL 106  
 XX  
 XX RESULT 23  
 XX AAR50228  
 XX ID AAR50228 standard; protein; 254 AA.  
 XX  
 XX AC AAR50228;  
 XX  
 XX DT 27-AUG-2003 (revised)  
 XX DT 25-MAR-2003 (revised)  
 XX DT 09-OCT-1994 (first entry)  
 XX  
 XX DB Sequence of a fragment of Group A streptococcal surface protein M5 and a  
 XX carrier of the COOH-terminal portion of M5.  
 XX  
 XX KW B subunit; labile toxin; M protein; fusion protein; antigen;  
 KW Group A streptococci; rheumatic fever; pharyngitis.  
 XX  
 XX OS Streptococcus sp.  
 XX

```
PN WO9406465-A1.
XX
XX 31-MAR-1994.
PD
XX 15-SEP-1993; 93WO-US008704.
XX
XX 16-SEP-1992; 92US-00945860.
XX
XX (UYTE-) UNIV TENNESSEE RES CORP.
PA
XX Dale JB;
XX
XX WPI; 1994-118162/14.
DR
XX N-PSDB; AAQ45160.
XX
XX New recombinant hybrid streptococcal M protein antigen(s) - which elicit
PT opsonic antibodies without eliciting cross-reactive antibodies to
PT mammalian heart tissue.
XX
XX Disclosure; Fig 4; 45pp; English.
XX
XX The surface M protein of Group A streptococci is the major virulence
CC factor and protective antigen of these organisms. However, there are a
CC tremendous number of M protein serotypes. The invention provides
CC recombinant M protein antigens comprising a gene encoding a carrier
CC protein and an NH2 or COOH terminal M protein fragment carrying one or
CC more epitopes. The carrier may be the B subunit of E.coli labile toxin
CC (LT-B) or the C-repeat portion of a streptococcal M protein. For example,
CC AAQ45160/R50229 comprises an antigen which is an M5 hapten fragment of 16
CC AAs joined by a BamHI restriction site to a carrier which is the COOH-
CC terminal half of M5. The carrier includes 2.5 C-repeats, which each
CC commence with the tetrapeptide NKIS. A linker could be inserted at the
CC BamHI site. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
XX Sequence 254 AA;
SQ
Query Match 51.8%; Score 71; DB 2; Length 254;
Best Local Similarity 75.0%; Pred. No. 1.2;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 9 ASREAKQVEKALEQLEDKV 28
DB 105 ASREAKQVEKALEEANSKL 124
RESULT 24
AAR50229
ID AAR50229 standard; protein; 284 AA.
XX
XX AAR50229;
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 09-OCT-1994 (first entry)
XX
XX Sequence of fragments of Group A streptococcal surface protein M5 and a
DE carrier of the COOH-terminal portion of M5.
XX
XX B subunit; labile toxin; M protein; fusion protein; antigen;
KM Group A streptococci; rheumatic fever; pharyngitis.
XX
XX Streptococcus sp.
OS
XX WO9406465-A1.
XX
XX 31-MAR-1994.
PD
XX 15-SEP-1993; 93WO-US008704.
XX
XX 16-SEP-1992; 92US-00945860.
XX
XX (UYTE-) UNIV TENNESSEE RES CORP.
PA
```

```
XX Dale JB;
XX
XX WPI; 1994-118162/14.
DR
XX N-PSDB; AAQ45161.
XX
XX New recombinant hybrid streptococcal M protein antigen(s) - which elicit
PT opsonic antibodies without eliciting cross-reactive antibodies to
PT mammalian heart tissue.
XX
XX Disclosure; Fig 5; 45pp; English.
XX
XX The surface M protein of Group A streptococci is the major virulence
CC factor and protective antigen of these organisms. However, there are a
CC tremendous number of M protein serotypes. The invention provides
CC recombinant M protein antigens comprising a gene encoding a carrier
CC protein and an NH2 or COOH terminal M protein fragment carrying one or
CC more epitopes. The carrier may be the B subunit of E.coli labile toxin
CC (LT-B) or the C-repeat portion of a streptococcal M protein. For example,
CC AAQ45161/R50229 comprises three segments of M5 designated A, B and C. The
CC C segment is joined by a BamHI restriction site to a carrier, which is the
CC COOH-terminal half of M5. The carrier includes 2.5 C-repeats, which each
CC commence with the tetrapeptide NKIS. A linker could be inserted at the
CC BamHI site. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
XX Sequence 284 AA;
SQ
Query Match 51.8%; Score 71; DB 2; Length 284;
Best Local Similarity 75.0%; Pred. No. 1.3;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 9 ASREAKQVEKALEQLEDKV 28
DB 135 ASREAKQVEKALEEANSKL 154
RESULT 25
AAR50996
ID AAR50996 standard; protein; 305 AA.
XX
XX AAR50996;
XX
XX 25-MAR-2003 (revised)
DT 02-NOV-1994 (first entry)
XX
XX Recombinant M24-M5-M6-M19 C-term variant.
XX
XX primer; PCR; amplify; polymerase chain reaction; construct; hybrid;
KM M protein; M24; M5; M6; M19; subunit; tetraivalent protein; emm gene;
KM tandem; PKK223.3; protective epitope; tissue-cross-reactive epitope;
KM restriction enzyme site; multivalent M protein; immunisation; group A;
KM streptococci; rheumatic fever; rheumatic heart disease; humoral;
KM antibody; heart tissue; antigen; serotype; mucosal.
XX
XX Synthetic.
OS
XX Key
XX location/Qualifiers
XX 1..18
FH Peptide /label= M24
FT 19..35
FT /label= M5
FT 36..52
FT /label= M6
FT 53..69
FT /label= M19
FT 70..305
FT /label= M5_C-terminal_fragment
XX
XX WO9406421-A1.
XX
XX 31-MAR-1994.
PD
XX
```



PF 15-SEP-1993; 93WO-US008703.  
XX  
XX  
PR 16-SEP-1992; 92US-00945954.  
XX  
PA (UYTE-) UNIV TENNESSEE RES CORP.  
XX  
PI Dale JB, Lederer JW,  
XX  
XX WPI; 1994-118122/14.  
DR N-PSDB; AA045218.  
XX  
XX New immunogenic hybrid proteins derives from streptococcal M proteins -  
PT induces opsonic antibodies, for protective immunisation against multiple  
PT group A streptococci serotypes.  
XX  
XX  
PS Disclosure; Fig 8; 67pp; English.  
XX  
XX The sequences given in AAR50992-1001 represent hybrid M proteins which  
CC contain the M24-M5-M6 and/or M19 subunits. These multivalent proteins  
CC were constructed using fragments of the 5' regions of emm genes that were  
CC amplified by PCR, ligated in tandem and expressed in pKK223.3. The  
CC amplified regions pref. encode protective and not tissue-cross-reactive  
CC epitopes, which can then be linked into one protein molecule. The  
CC recombinant hybrid protein may contain 113 N-terminal amino acids of M24,  
CC 58 amino acids of M5, 35 from M6 and 35 from M19. Each section is linked  
CC by 2 amino acids specified by the respective restriction enzyme sites  
CC that were synthesised into the primers used to specify the PCR product.  
CC Multivalent M proteins such as this may be used for protective  
CC immunisation against group A streptococci, which esp. cause rheumatic  
CC fever and rheumatic heart disease. Humoral antibodies raised against  
CC these proteins do not react with heart tissue antigens but are effective  
CC against many different serotypes. The multivalent proteins may also  
CC include sequences which induce mucosal antibodies and do not require  
CC coupling to an immunogenic carrier. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX  
SQ Sequence 305 AA;  
XX  
Query Match 51.8%; Score 71; DB 2; Length 305;  
Best Local Similarity 75.0%; Pred. No. 1.4;  
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ASREAKKQVEKALBQLEDKV 28  
DB 156 ASREAKKQVEKALBQLEANSKL 175  
XX  
RESULT 26  
AAB03117  
ID AAB03117 standard; protein; 305 AA.  
XX  
XX AAB03117;  
AC  
XX  
DT 10-OCT-2000 (first entry)  
XX  
XX S. Pyogenes hybrid M protein (M24-M5-M6-M19), SEQ ID NO:10.  
DE  
XX  
XX Multivalent hybrid M protein; group A streptococcus; serotype;  
KM immunogenic; sero-specific antibody; streptococcal infection;  
KM cross reactivity; vaccine; acute rheumatic fever; ARF;  
KM rheumatic heart disease; streptococcal pharyngitis; strep throat;  
KM pneumonia.  
XX  
XX Streptococcus pyogenes.  
OS Synthetic.  
XX  
XX US6063386-A.  
PN  
XX 16-MAY-2000.  
PD  
XX 15-SEP-1997; 97US-00937271.  
PF  
XX 16-SEP-1992; 92US-00945954.  
PR

XX  
XX (UYTE-) UNIV TENNESSEE RES CORP.  
PA  
XX  
XX Lederer JW, Dale JB;  
PI  
XX  
XX WPI; 2000-364475/31.  
DR  
XX  
XX N-PSDB; AAA57897.  
DR  
XX  
XX New immunogenic recombinant hybrid M protein comprising amino-terminal  
PT peptide fragments of streptococcal M protein useful as vaccine against  
PT rheumatic fever and infections leading to rheumatic fever.  
XX  
XX  
PS Disclosure; Fig 8A-B; 62pp; English.  
XX  
XX The invention relates to multivalent immunogenic hybrid group A  
CC streptococcal M proteins comprising N-terminal peptide fragments of M  
CC proteins that elicit opsonic antibodies against multiple serotypes of  
CC group A streptococci (e.g., Streptococcus pyogenes). The antibodies  
CC generated using the hybrid proteins are against one or more M protein  
CC serotypes, where at least one of the serotypes is M1, M3, M18 or M19. The  
CC invention also encompasses a recombinant DNA molecule comprising a  
CC nucleotide sequence that encodes a multivalent hybrid M protein; and a  
CC method for immunising a mammal against streptococcal infections,  
CC comprising administering an immunogenic multivalent hybrid M protein to  
CC the mammal. The multivalent hybrid M proteins are useful for eliciting  
CC opsonic or protective antibodies to the M proteins of different serotypes  
CC of group A streptococci, and may therefore be used as vaccines to protect  
CC against and control infection by type A streptococci. Type A streptococci  
CC are not only responsible for streptococcal pharyngitis (strep throat),  
CC forms of pneumonia and a condition resembling toxic shock, but are also  
CC involved in the development of acute rheumatic fever (ARF) and rheumatic  
CC heart disease. In a patient with ARF, antibodies formed during a group A  
CC streptococcal infection are also cross-reactive with heart tissue, which  
CC indicates that the streptococci and host tissue contain similar antigenic  
CC motifs. The new multivalent vaccines are capable of raising sero-specific  
CC antibodies against various serotypes of group A streptococci which are  
CC not cross-reactive with human heart tissue. Sequences AAB03113-B03117,  
CC AAB03119-B03121 and AAB03123-B03124 represent multivalent hybrid  
CC Streptococcus pyogenes M proteins generated in the disclosure of the  
CC invention  
XX  
SQ Sequence 305 AA;  
XX  
Query Match 51.8%; Score 71; DB 3; Length 305;  
Best Local Similarity 75.0%; Pred. No. 1.4;  
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ASREAKKQVEKALBQLEDKV 28  
DB 156 ASREAKKQVEKALBQLEANSKL 175  
XX  
RESULT 27  
ADP49327  
ID ADP49327 standard; protein; 539 AA.  
XX  
XX ADP49327;  
AC  
XX  
DT 26-AUG-2004 (first entry)  
XX  
XX S pyogenes M21 protein.  
DE  
XX  
XX antiporiatic; vaccine; psoriasis; alpha-helix coiled-coil rod homologue;  
KM HCR; polymorphism; M21.  
KM  
XX Streptococcus pyogenes.  
OS  
XX  
XX WO2004046380-A1.  
PN  
XX 03-JUN-2004.  
PD  
XX 19-NOV-2003; 2003WO-GB005011.  
PF  
XX

PR 20-NOV-2002; 2002GB-00027026.  
 XX (MOLE-) MOLECULAR SKINCARE LTD.  
 PA  
 PI Cork MJ, Ward SJ, Tazi-Ahmini R;  
 XX  
 DR WPI; 2004-449755/42.  
 XX  
 PT Diagnosing or detecting polymorphism in alpha-helix coiled-coil rod  
 PT homolog (HCR) gene in human, by determining sequence of HCR gene in  
 PT polymorphic position and determining status of human by reference to  
 PT polymorphism in HCR.  
 XX  
 PS Claim 14; Page 48-50; 55pp; English.  
 XX  
 CC The present invention relates to a method of diagnosing or detecting a  
 CC polymorphism in the alpha-helix coiled-coil rod homologue (HCR) gene in a  
 CC human, which involves determining the sequence of the human HCR gene in  
 CC at least one polymorphic position of HCR and determining the status of  
 CC the human by reference to the polymorphism in HCR. The method is useful  
 CC for diagnosing or detecting a polymorphism in HCR gene in a human, which  
 CC can in turn be used for developing a diagnostic assay for susceptibility  
 CC and/or predisposition to guttate psoriasis or chronic plaque psoriasis.  
 CC Polymorphisms diagnosed or detected using the method are useful for  
 CC developing a medicament for treating guttate psoriasis or psoriasis  
 CC vulgaris, where the medicament interferes with the antigenic properties  
 CC of the polymorphic protein. The present sequence is the Streptococcus  
 CC pyogenes M21 protein, which can be used as a target for the medicament  
 CC described in the invention.  
 CC  
 SQ Sequence 539 AA;  
 Query Match 51.8%; Score 71; DB 8; Length 539;  
 Best Local Similarity 75.0%; Pred. No. 2.7;  
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 9 ASREAKKQVEKALEQLEDKV 28  
 DB 390 ASREAKKQVEKALEBANSKL 409  
 RESULT 28  
 AAR97453  
 ID AAR97453 standard; peptide; 28 AA.  
 XX  
 AC AAR97453;  
 XX  
 DT 04-DEC-1996 (first entry)  
 XX  
 DE Chimaeric peptide av85 contg. C. elegans unc-15 paramyosin peptide.  
 XX  
 KW Paramyosin protein; peptide; unc-15; chimaeric; chimeric; B-cell;  
 KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
 KW mapping; opsonic antibody; vaccine; immunotherapy; diagnosis.  
 XX  
 OS Synthetic.  
 XX  
 PN MO9611944-A1.  
 XX  
 PD 25-APR-1996.  
 XX  
 PF 16-OCT-1995; 95WO-AU000681.  
 XX  
 PR 14-OCT-1994; 94AU-00008851.  
 XX  
 PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.  
 PA (UYME) UNIV MELBOURNE.  
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
 PA (CSLC-) CSL LTD.  
 XX  
 PI Cooper JA, Relf WA, Good MF, Saul AJ;

XX  
 DR WPI; 1996-221939/22.  
 XX  
 PT New chimeric peptide(s) including a conformational epitope - inserted  
 PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
 XX  
 PS Example 18; Page 44; 99pp; English.  
 XX  
 CC The present peptide is a chimaeric peptide (CP), contg. a C. elegans unc-  
 CC 15 paramyosin peptide. The CP comprises a B-cell conformational epitope  
 CC from within unc-15, inserted into a 2nd peptide, pref. an alpha-helical  
 CC coil based on the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide  
 CC has a similar conformation, enabling the epitope to be presented in an  
 CC immunologically active conformation. The CP can be used in a novel  
 CC detection/mapping process, e.g. to determine the min. epitope required to  
 CC induce opsonic antibodies (Ab), and in vaccines against C. elegans. Ab  
 CC raised against the CP can be used for immunotherapy and diagnosis, while  
 CC the CP can be used diagnostically to detect Ab.  
 CC  
 SQ Sequence 28 AA;  
 Query Match 50.0%; Score 68.5; DB 2; Length 28;  
 Best Local Similarity 55.2%; Pred. No. 0.21;  
 Matches 16; Conservative 4; Mismatches 2; Indels 7; Gaps 1;  
 QY 1 KQAEKVKASREAKKQVEKALEQLEDKV 29  
 DB 4 KQAEKVK-----KQLEDKVEBELQDKV 25  
 RESULT 29  
 AAR97439  
 ID AAR97439 standard; peptide; 29 AA.  
 XX  
 AC AAR97439;  
 XX  
 DT 02-DEC-1996 (first entry)  
 XX  
 DE Chimaeric peptide bd11 contg. C. elegans unc-15 paramyosin peptide.  
 XX  
 KW Paramyosin protein; peptide; unc-15; chimaeric; chimeric; B-cell;  
 KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
 KW mapping; opsonic antibody; vaccine; immunotherapy; diagnosis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 8..22  
 FT /note= "unc-15 conformational B-cell epitope"  
 XX  
 PN MO9611944-A1.  
 XX  
 PD 25-APR-1996.  
 XX  
 PF 16-OCT-1995; 95WO-AU000681.  
 XX  
 PR 14-OCT-1994; 94AU-00008851.  
 XX  
 PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.  
 PA (UYME) UNIV MELBOURNE.  
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
 PA (CSLC-) CSL LTD.  
 XX  
 PI Cooper JA, Relf WA, Good MF, Saul AJ;  
 XX  
 DR WPI; 1996-221939/22.  
 XX  
 PT New chimeric peptide(s) including a conformational epitope - inserted  
 PT into a peptide having similar native conformation, useful in vaccines and  
 PT for determin. of minimal epitope(s) or for mapping amphipathic helices.

XX Example 18; Page 44; 99pp; English.  
XX  
CC The present peptide is a chimeric peptide (CP) contg. a C. elegans unc-  
XX 15 paramecin peptide. The CP comprises a B-cell conformational epitope  
CC from within unc-15, inserted into a 2nd peptide, pref. an alpha-helical  
CC coil based on the GCN4 leucine zipper peptide (AAR7395). The 2nd peptide  
CC has a similar conformation, enabling the epitope to be presented in an  
CC immunologically active conformation. The CP can be used in a novel  
CC detection/mapping process, e.g. to determine the min. epitope required to  
CC induce opsonic antibodies (Ab), and in vaccines against C. elegans. Ab  
CC raised against the CP can be used for immunotherapy and diagnosis, while  
CC the CP can be used diagnostically to detect Ab  
XX  
SQ Sequence 29 AA;  
Query Match 48.9%; Score 67; DB 2; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.33;  
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
OY 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 29  
Db 1 KOAEDKVKMAQDTADRLTEKLNQLEDKVK 29  
RESULT 30  
ABU39221  
ID ABU39221 standard; protein; 389 AA.  
XX  
AC ABU39221;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by prokaryotic essential gene #24748.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Pasteurella multocida.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR N-PSDB; ACA43091.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 67145; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway of  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than S. aureus, S. typhimurium,  
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 389 AA;  
Query Match 48.9%; Score 67; DB 6; Length 389;  
Best Local Similarity 51.6%; Pred. No. 5.7;  
Matches 16; Conservative 5; Mismatches 8; Indels 2; Gaps 1;  
OY 1 KOAED--KVKASREAKKQVEKALEQLEDKVK 29  
Db 175 KOAEEBAKAKAEEBAKKAERAKKAERAKAK 205  
RESULT 31  
ADL99394  
ID ADL99394 standard; protein; 284 AA.  
XX  
AC ADL99394;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Nanostructure assembly protein #45.  
XX  
KW peptide nucleic acid; PNA; nanostructure.  
XX  
OS Synthetic.  
XX  
PN US2003215903-A1.  
XX  
PD 20-NOV-2003.  
XX  
PF 21-FEB-2003; 2003US-00370685.  
XX  
PR 21-FEB-2002; 2002US-00080608.  
XX  
PA (HYMA/) HYMAN P L.  
PA (GOLD/) GOLDBERG E B.  
XX  
PI Hyman PL, Goldberg EB;  
XX  
DR WPI; 2004-021840/02.  
XX  
XX Staged assembly of a nanostructure containing peptide nucleic acid  
PT assembly units comprising contacting a nanostructure intermediate with an  
PT assembly unit comprising different joining elements, and removing unbound  
PT assembly units.  
XX  
PS Disclosure; Page 40-41; 73pp; English.  
XX  
XX The invention relates to staged assembly of a nanostructure comprising:  
CC (a) contacting a nanostructure intermediate comprising at least one



DR WPI; 1996-221939/22.  
XX  
XX New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
PS  
PS Example 18; Page 45; 99pp; English.  
XX  
CC The present peptide is a chimeric peptide (CP), contg. a C. elegans unc-  
CC 15 parmyosin peptide. The CP comprises a B-cell conformational epitope  
CC from within unc-15, inserted into a 2nd peptide, pref. an alpha-helical  
CC coil based on the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide  
CC has a similar conformation, enabling the epitope to be presented in an  
CC immunologically active conformation. The CP can be used in a novel  
CC detection/mapping process, e.g. to determine the min. epitope required to  
CC induce opsonic antibodies (Ab), and in vaccines against C. elegans. Ab  
CC the CP can be used diagnostically to detect Ab  
XX  
XX Sequence 29 AA;  
SQ  
Query Match 47.4%; Score 65; DB 2; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.58; Mismatches 9; Indels 0; Gaps 0;  
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
OY 1 KOAEDKVASREAKKQVEKALEQLEDEKVK 29  
DB 1 KQLEEKVKQAEKELNIQKQKQLQDQKVK 29  
RESULT 34  
AAR43343  
ID AAR43343 standard; peptide; 20 AA.  
XX  
XX AAR43343;  
XX  
XX 25-MAR-2003 (revised)  
DT 21-MAY-1994 (first entry)  
XX  
XX B cell epitope from Streptococcus M protein.  
DE  
XX Group A beta haemolytic streptococci; antibody; human heart tissue;  
KW vaccine; humoral immunity; diagnosis.  
XX  
XX Synthetic.  
OS  
XX MO9321220-A1.  
PN  
XX 28-OCT-1993.  
PD  
XX 30-MAR-1993; 93WO-AU000131.  
PF  
XX 08-APR-1992; 92AU-00001800.  
PR  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA  
XX  
PI Good MF, Pruksakorn S;  
XX  
XX WPI; 1993-351655/44.  
DR  
XX Synthetic peptide - comprises at least one B-cell epitope from  
PT streptococcal M-protein, useful in vaccine for streptococcal infections.  
XX  
XX Claim 7; Page 38; 57pp; English.  
PS  
XX The peptide comprises an N-terminal sequence derived from the conserved  
CC region of the M protein of group A beta-haemolytic streptococci (residues  
CC 337-492 of the type 5 M-protein). The peptide shown, peptide 145  
CC comprises at least 1 B cell epitope, where an antibody reactive to the B  
CC cell epitope is only minimally reactive to human heart tissue. The  
CC peptide is useful in a vaccine for inducing humoral immunity against  
CC streptococcal infections. Antibodies to the peptide are useful in  
CC diagnosis of such infections. (Updated on 25-MAR-2003 to correct PN

CC field.)  
XX  
XX Sequence 20 AA;  
SQ  
Query Match 46.7%; Score 64; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.51; Mismatches 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 9 ASREAKKQVEKALE 22  
DB 7 ASREAKKQVEKALE 20  
RESULT 35  
AAR97390  
ID AAR97390 standard; peptide; 20 AA.  
XX  
XX AAR97390;  
AC  
XX  
XX 27-AUG-2003 (revised)  
DT 02-DEC-1996 (first entry)  
XX  
XX Streptococcal M protein peptide p145, for use in chimeric peptide.  
DE  
XX Streptococcal; M protein; peptide; p145; chimeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; group A Streptococci; immunotherapy;  
KW diagnosis.  
XX  
XX Streptococcus sp.  
OS  
XX  
XX MO9611944-A1.  
PN  
XX  
XX 25-APR-1996.  
PD  
XX  
XX 16-OCT-1995; 95WO-AU000681.  
PF  
XX 14-OCT-1994; 94AU-00008851.  
PR  
XX  
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (CSTR ) COMMONWEALTH SCI & IND RES ORG.  
PA (UYME ) UNIV MELBOURNE.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSLC-) CSL LTD.  
XX  
XX Cooper JA, Relf WA, Good MF, Saul AJ;  
PI  
XX  
XX WPI; 1996-221939/22.  
DR  
XX  
XX New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determin. of minimal epitope(s) or for mapping amphipathic helices.  
PS  
PS Claim 4; Page 80; 99pp; English.  
XX  
XX The present peptide is the Streptococcal M protein peptide p145  
CC (Pruksakorn et al, J. Immunol. 149: 2729-2735 (1992)), used in the  
CC construction of a novel chimeric peptide (CP). The CP comprises a B-cell  
CC conformational epitope from within the present peptide, inserted into a  
CC 2nd peptide, pref. an alpha-helical coil based on the GCN4 leucine zipper  
CC peptide (AAR97395). The 2nd peptide has a similar conformation, enabling  
CC the epitope to be presented in an immunologically active conformation.  
CC The CP can be used in a novel detection/mapping process, e.g. to  
CC determine the min. epitope required to induce opsonic antibodies (Ab),  
CC and in vaccines against sp. A Streptococci. Ab raised against the CP can  
CC be used for immunotherapy and diagnosis, while the CP can be used  
CC diagnostically to detect Ab. (Updated on 27-AUG-2003 to correct OS  
CC field.)  
XX  
XX Sequence 20 AA;  
SQ  
Query Match 46.7%; Score 64; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALE 22  
Db 7 ASREAKKQVEKALE 20

RESULT 36  
AAW69281  
ID AAW69281 standard; peptide; 20 AA.

AC AAW69281;  
DT 29-OCT-1998 (first entry)

DE Streptococcus group A protein fragment.

XX Acryloylated peptide polymer; immune response; peptide epitope;  
KW synthetic vaccine; enzymatically cleavable site.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "linked to acryloylated peptide polymer"

PN MO9834968-A1.

PD 13-AUG-1998.

PF 10-FEB-1998; 98WO-AU000076.

PR 11-FEB-1997; 97AU-00005071.

PR 03-OCT-1997; 97CA-02217321.

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (UYME) UNIV MELBOURNE.

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

PA (CSLC-) CSL LTD.

PI Jackson DC, Obrien-Simpson NM, Brown LE, Zeng W, Ede NJ;

PI Brandt ER, Good MF;

DR WPI, 1998-447177/38.

XX Acryloylated peptide polymers - useful for synthetic vaccine technology,  
PT for raising an immune response to peptide epitope and as diagnostic tool.

PS Example 1; Page 20; 77pp; English.

XX This sequence represents a fragment of a Streptococcus group A protein.

CC This sequence was used to test the acryloylated peptide polymer of the  
CC invention. The peptide polymers are used to raise an immune response to a  
CC peptide epitope (such as this sequence), and also as diagnostic tools.

CC Polymers (molecular wt. >600 kDa.) can be prepared with virtually any  
CC number of the same or different epitopes by a method that allows  
CC purification of the individual determinants, avoids errors inherent in  
CC long sequential syntheses in which protected peptide fragments are not  
CC used, thus avoiding solubility and purification problems. Multiple copies  
CC of many different peptide epitopes may be incorporated into a single  
CC polymeric structure to allow utilisation of the range of T cell epitopes  
CC required for outbred populations in conjunction with epitopes  
CC representing different pathogenic serotypes, thus making them a  
CC significant advance in synthetic vaccine technology

XX Sequence 20 AA;

Query Match 46.7%; Score 64; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.51; 0; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;

QY 9 ASREAKKQVEKALE 22  
Db 7 ASREAKKQVEKALE 20

RESULT 37  
AAR97460  
ID AAR97460 standard; peptide; 29 AA.

AC AAR97460;

DT 04-DEC-1996 (first entry)

DE Chimeric peptide bd7 contg. C. elegans unc-15 paramyosin peptide.

XX Paramyosin protein; peptide; unc-15; chimeric; chimeric; B-cell;  
KW conformational epitope; alpha-helix; GCN4; leucine zipper; detection;  
KW mapping; opsonic antibody; vaccine; immunotherapy; diagnosis.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 11..21 /note= "unc-15 conformational B-cell epitope"

PN MO9611944-A1.

PD 25-APR-1996.

PF 16-OCT-1995; 95WO-AU0000681.

PR 14-OCT-1994; 94AU-00008851.

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (UYME) UNIV MELBOURNE.

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.

PA (CSLC-) CSL LTD.

PI Cooper JA, Relf WA, Good MF, Saul AJ;

DR WPI, 1996-221939/22.

XX New chimeric peptide(s) including a conformational epitope - inserted  
PT into a peptide having similar native conformation, useful in vaccines and  
PT for determ. of minimal epitope(s) or for mapping amphipathic helices.

PS Example 18; Page 45; 99pp; English.

XX The present peptide is a chimeric peptide (CP), contg. a C. elegans unc-  
CC 15 paramyosin peptide. The CP comprises a B-cell conformational epitope  
CC from within unc-15, inserted into a 2nd peptide (AAR97395). The 2nd peptide  
CC coil based on the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide  
CC has a similar conformation, enabling the epitope to be presented in an  
CC immunologically active conformation. The CP can be used in a novel  
CC detection/mapping process, e.g. to determine the min. epitope required to  
CC induce opsonic antibodies (Ab), and in vaccines against C. elegans. Ab  
CC raised against the CP can be used for immunotherapy and diagnosis, while  
CC the CP can be used diagnostically to detect Ab

XX Sequence 29 AA;

Query Match 46.7%; Score 64; DB 2; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.77;  
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 KOAEKXVYASREAKKQVEKALEQLEDKVK 29  
Db 1 KQLEBKXVQATEKLNIGKQKQLQDKVK 29

RESULT 38



```

PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2003-895291/82.
DR N-PSDB; ADF00933.
XX
PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
PS Disclosure; SEQ ID NO 5390; 870pp; English.
XX
CC The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
SQ Sequence 361 AA;

```

Query Match 46.7%; Score 64; DB 7; Length 361;  
Best Local Similarity 48.3%; Pred. No. 12;  
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

```

OY 1 KOAEDKVKASREAKQVEKALEQLQEDKVK 29
   :|||:|||||:|:|:|:|:|:|:|:|:|
DB 165 EOADKAKAOEAERAKKQAEIQAQKAKAEAK 193

```

Search completed: June 13, 2005, 20:40:51  
Job time : 159.488 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 20:37:05 ; Search time 40.4651 Seconds  
(without alignments)  
53.498 Million cell updates/sec

Title: US-10-706-275-2

Perfect score: 137

Sequence: 1 KOAEDKVKASREAKQVEKALEQLEDKVK 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	80.3	28	3	US-08-817-811-18
2	96	70.1	28	3	US-08-817-811-17
3	81	59.1	28	3	US-08-817-811-16
4	78	56.9	28	3	US-08-817-811-15
5	72	52.6	440	4	US-08-302-756E-35
6	72	52.6	443	2	US-08-795-475-6
7	72	52.6	443	4	US-08-325-278B-6
8	71	51.8	28	3	US-08-817-811-12
9	71	51.8	236	4	US-08-937-271-11
10	71	51.8	254	3	US-08-914-479A-4
11	71	51.8	284	4	US-08-914-479A-6
12	71	51.8	305	3	US-08-937-271-10
13	68.5	50.0	28	3	US-08-817-811-66
14	67	48.9	29	3	US-08-817-811-52
15	65.5	47.8	29	3	US-08-817-811-15
16	64	46.7	29	3	US-08-817-811-74
17	64	46.7	20	3	US-08-817-811-1
18	64	46.7	29	3	US-08-817-811-73
19	64	46.7	72	1	US-08-182-175A-87
20	64	46.7	72	5	PCT-US92-06412-87
21	64	46.7	361	4	US-09-543-681A-5390
22	63	46.0	29	3	US-08-817-811-79
23	63	46.0	107	1	US-08-182-175A-105
24	63	46.0	107	1	US-08-474-633A-92
25	63	46.0	107	4	US-08-823-771-92
26	63	46.0	107	5	PCT-US92-06412-105
27	62	45.3	28	3	US-08-817-811-13

28	62	45.3	29	3	US-08-817-811-78	Sequence 78, Appl
29	62	45.3	77	1	US-08-182-175A-57	Sequence 57, Appl
30	62	45.3	77	1	US-08-474-633A-75	Sequence 75, Appl
31	62	45.3	77	4	US-08-823-771-75	Sequence 75, Appl
32	62	45.3	77	5	PCT-US92-06412-57	Sequence 57, Appl
33	61	44.5	28	1	US-08-182-175A-49	Sequence 49, Appl
34	61	44.5	28	4	US-08-474-633A-58	Sequence 58, Appl
35	61	44.5	28	1	US-08-823-771-58	Sequence 58, Appl
36	61	44.5	28	5	PCT-US92-06412-49	Sequence 49, Appl
37	61	44.5	29	3	US-08-817-811-71	Sequence 71, Appl
38	61	44.5	29	3	US-08-817-811-72	Sequence 72, Appl
39	61	44.5	29	3	US-08-817-811-75	Sequence 75, Appl
40	61	44.5	29	3	US-08-817-811-89	Sequence 89, Appl
41	61	44.5	37	1	US-08-182-175A-85	Sequence 85, Appl
42	61	44.5	37	1	US-08-182-175A-97	Sequence 97, Appl
43	61	44.5	37	1	US-08-474-633A-85	Sequence 85, Appl
44	61	44.5	37	1	US-08-474-633A-86	Sequence 86, Appl
45	61	44.5	37	4	US-08-823-771-85	Sequence 85, Appl

## ALIGNMENTS

RESULT 1  
US-08-817-811-18  
; Sequence 18, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Reif, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; TITLE OF INVENTION: COMPRISING SAME  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817, 811  
; FILING DATE: 14-APR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 96/11944  
; FILING DATE: 25-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: PIRC:005  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-08-817-811-18

Query Match 80.3%; Score 110; DB 3; Length 28;  
Best Local Similarity 85.2%; Pred. No. 1.6e-06;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KOAEDKVKASREAKKQVEKALEOLEDK 27  
|||:|||||:|||||  
Db 2 KOAEDKVDASREAKKQVKVKQKQLEDK 28

## RESULT 2

US-08-817-811-17  
; Sequence 17, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Relif, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; TITLE OF INVENTION: COMPRISING SAME  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817, 811  
; FILING DATE: 14-APR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 96/11944  
; FILING DATE: 25-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: FBRC:005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-817-811-17

Query Match 70.1%; Score 96; DB 3; Length 28;  
Best Local Similarity 76.9%; Pred. No. 6.4e-05;  
Matches 20; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KOAEDKVKASREAKKQVEKALEOLEDK 26  
|||:|||||:|||||  
Db 3 KOAEDKVDASREAKKQVKVKQKQLEDK 28

## RESULT 3

US-08-817-811-16  
; Sequence 16, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Relif, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; TITLE OF INVENTION: COMPRISING SAME  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817, 811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-16

Query Match 59.1%; Score 81; DB 3; Length 28;  
Best Local Similarity 68.0%; Pred. No. 0.0035;  
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KOAEDKVKASREAKKQVEKALEOLEDK 25  
|||:|||||:|||||  
Db 4 KOAEDKVDASREAKKQVKVKQKQLEDK 28

## RESULT 4

US-08-817-811-67  
; Sequence 67, Application US/08817811  
; Patent No. 6174528  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Juan A.  
; APPLICANT: Relif, Wendy A.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Saul, Allan J.  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
; TITLE OF INVENTION: COMPRISING SAME  
; NUMBER OF SEQUENCES: 97  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817, 811  
; FILING DATE: 14-APR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 96/11944  
; FILING DATE: 25-APR-1996

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Highlander, Steven L.
/ REGISTRATION NUMBER: 37,642
/ REFERENCE/DOCKET NUMBER: FBRC:005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512/418-3000
/ TELEFAX: 512/474-7577
/ INFORMATION FOR SEQ ID NO: 67:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 28 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
US-08-817-811-67

Query Match          56.9%; Score 78; DB 3; Length 28;
Best Local Similarity 64.0%; Pred. No. 0.0077;
Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Cy 1 KOAEDKVKASREAKKQVEKALEQLE 25
    |||||:|||||:|||||:
    4 KOAEDDLASREAKKQLQDKVKNQLE 28

RESULT 5
US-08-302-756E-35
/ Sequence 35, Application US/08302756E
/ Patent No. 6737521
/ GENERAL INFORMATION:
/ APPLICANT: FISCHETTI, Vincent A.
/ APPLICANT: POZZI, Gianni
/ TITLE OF INVENTION: DELIVERY AND EXPRESSION OF A HYBRID SURFACE PROTEIN ON
/ FILE REFERENCE: 016921-076
/ CURRENT APPLICATION NUMBER: US/08/302,756E
/ PRIOR FILING DATE: 1995-03-07
/ PRIOR APPLICATION NUMBER: US 07/522,440
/ PRIOR FILING DATE: 1990-05-11
/ PRIOR APPLICATION NUMBER: US 07/742,199
/ PRIOR FILING DATE: 1991-08-05
/ PRIOR APPLICATION NUMBER: US 07/814,823
/ PRIOR FILING DATE: 1991-12-23
/ PRIOR APPLICATION NUMBER: US 07/851,082
/ PRIOR FILING DATE: 1992-03-13
/ PRIOR APPLICATION NUMBER: PCT/US93/02355
/ PRIOR FILING DATE: 1993-03-12
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 35
/ LENGTH: 440
/ TYPE: PRT
/ ORGANISM: S. pyogenes
US-08-302-756E-35

Query Match          52.6%; Score 72; DB 4; Length 440;
Best Local Similarity 45.2%; Pred. No. 0.64;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKQVEKALEQLEDRV 28
    |||||:|||||:|||||:
    270 DKVKEKQISDASRQGLRDLASREAKQVEKALEANSKTL 311

Db 270 DKVKEKQISDASRQGLRDLASREAKQVEKALEANSKTL 311

RESULT 6
US-08-795-475-6
/ Sequence 6, Application US/08795475
/ Patent No. 5965390
/ GENERAL INFORMATION:
/ APPLICANT: Bjorck, Lars
/ APPLICANT: Sjodring, Ulf
/ TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
/ NUMBER OF SEQUENCES: 14
```

```
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SEED AND BERRY LLP
/ STREET: 6300 Columbia Center, 701 Fifth Avenue
/ CITY: Seattle
/ STATE: Washington
/ COUNTRY: USA
/ ZIP: 98104-7092
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/795,475
/ FILING DATE: 11-FEB-1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McMaisters, David D.
/ REGISTRATION NUMBER: 33,963
/ REFERENCE/DOCKET NUMBER: 100084.402D1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 443 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-795-475-6

Query Match          52.6%; Score 72; DB 2; Length 443;
Best Local Similarity 45.2%; Pred. No. 0.65;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKQVEKALEQLEDRV 28
    |||||:|||||:|||||:
    272 DKVKEKQISDASRQGLRDLASREAKQVEKALEANSKTL 313

Db 272 DKVKEKQISDASRQGLRDLASREAKQVEKALEANSKTL 313

RESULT 7
US-08-325-278B-6
/ Sequence 6, Application US/08325278B
/ Patent No. 6822075
/ GENERAL INFORMATION:
/ APPLICANT: Bjorck, Lars
/ APPLICANT: Sjodring, Ulf
/ TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Seed IP Law Group
/ STREET: 701 Fifth Avenue Suite 6300
/ CITY: Seattle,
/ STATE: Washington
/ COUNTRY: USA
/ ZIP: 98104-7092
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/325,278B
/ FILING DATE: 26-Oct-1994
/ CLASSIFICATION: <unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Potter, Jane E. R.
/ REGISTRATION NUMBER: 33,332
/ REFERENCE/DOCKET NUMBER: 100084.402
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 6:
```

```

:      SEQUENCE CHARACTERISTICS:
:          LENGTH: 443 amino acids
:          TYPE: amino acid
:          TOPOLOGY: linear
:      MOLECULE TYPE: protein
:      SEQUENCE DESCRIPTION: SEQ ID NO: 6
:
IS-08-325-278B-6

```

Query Match	52.6%;	Score 72;	DB 4;	Length 443;
Beet Local Similarity	45.2%;	Pred. No. 0.65;		
Matches 19; Conservative	2;	Mismatches 3;	Indels 18;	Gaps 1;

```

QY      5  DKVK-----ASREAKKQVEKALEQLEDKV  28
          |||          |||||:|
DB      272 DKYKEKQISDASRQRLRDLSDASREAKKQVEKALEANSKL  313

```

RESULT 8  
US-08-817-811-12  
; Sequence 12, Application US/08817811

APPLICANT: Cooper, Juan A.  
APPLICANT: Relf, Wendy A.  
APPLICANT: Good, Michael F.

TITLE OF INVENTION: COMPRISING SAME  
 NUMBER OF SEQUENCES: 97  
 CORRESPONDENCE ADDRESS:

```

;
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
;

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS

```

APPLICATION NUMBER: US/08/017,811  
 FILING DATE: 14-APR-1997  
 CLASSIFICATION: 424

AFFILIATION NUMBER: NO 20/1137  
 FILING DATE: 25-APR-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Highlander, Steven L.

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577

SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:

US-06-817-811-12

Query Match

Best Local Similarity

Qy 1 KQAEDEVKASREAKKQVEKALEQLEDK 27  
Db 2 KQLEDEKVKOLEDEKVKOLEDEKVKOLEDEK 28

## RESULT 9

US-08-937-271-11  
; Sequence 11, Application US/08937271

APPLICANT: Dale, James B.  
APPLICANT: Lederer, James W.  
TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN

CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue

COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

FILED DATE: 10 SEP 1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Rosenman, Stephen J

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

SEQUENCE: UNKUNOINADIDK  
; LENGTH: 236 amino acid  
; TYPE: amino acid  
; TOPOLOGY: linear

Query Match	51.8%	Score 71;	DB 3;	Length 236
Best Local Similarity	75.0%	Pred. No. 0.44;		

QY 9 ASREAKKQVEKALEQLEDKV 28  
||| ||| :  
Db 87 ASREAKKQVEKALEANSKL 10

RESULT 10  
US-08-914-479A-4  
Sequence 4, Application US/08914479

```

: GENERAL INFORMATION:
: APPLICANT: Dale, James B.
: TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
: TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE

```

; CURRENT APPLICATION NUMBER: US/08/314,412  
 ; CURRENT FILING DATE: 1997-08-19  
 ; PRIOR APPLICATION NUMBER: 08/409,270  
 ; PRIOR FILING DATE: 995-03-23

```

; PRIOR FILING DATE: 1992-09-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```

```

; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: An antigen of M5 and a carrier of the

```

OTHER INFORMATION: COOH-terminal portion of M5

Page 4

US-08-914-479A-4

Query Match 51.8%; Score 71; DB 4; Length 254;  
Best Local Similarity 75.0%; Pred. No. 0.48;  
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALEBNSKL 28

DB 105 ASREAKKQVEKALEBNSKL 124

RESULT 11

US-08-914-479A-6

Sequence 6, Application US/08914479A

Patent No. 6419932

GENERAL INFORMATION:

APPLICANT: Dale, James B.

TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER

FILE REFERENCE: 481112.404C2

CURRENT APPLICATION NUMBER: US/08/914,479A

PRIOR FILING DATE: 1997-08-19

PRIOR APPLICATION NUMBER: 08/409,270

PRIOR FILING DATE: 1995-03-23

PRIOR APPLICATION NUMBER: 07/945,860

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 6

LENGTH: 284

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: An antigen of three fragments of M5 and a carrier

US-08-914-479A-6

Query Match

Best Local Similarity 75.0%; Pred. No. 0.53;

Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALEBNSKL 28

DB 135 ASREAKKQVEKALEBNSKL 154

RESULT 12

US-08-937-271-10

Sequence 10, Application US/08937271

Patent No. 6063386

GENERAL INFORMATION:

APPLICANT: Dale, James B.

TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/937,271

FILING DATE: 15-SEP-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Rosenman, Stephen J.

REGISTRATION NUMBER: 43,058

REFERENCE/DOCKET NUMBER: 481112.405C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 305 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-937-271-10

Query Match

Best Local Similarity 75.0%; Pred. No. 0.58;

Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALEBNSKL 28

DB 156 ASREAKKQVEKALEBNSKL 175

RESULT 13

US-08-817-811-66

Sequence 66, Application US/08817811

Patent No. 6174528

GENERAL INFORMATION:

APPLICANT: Cooper, Juan A.

APPLICANT: Reif, Wendy A.

APPLICANT: Good, Michael F.

APPLICANT: Saul, Allan J.

TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES

NUMBER OF SEQUENCES: 97

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White &amp; Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/817,811

FILING DATE: 14-APR-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO 96/11944

FILING DATE: 25-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/474-7577

TELEFAX: 512/418-3000

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 28 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-817-811-66

Query Match

Best Local Similarity 50.0%; Score 68.5; DB 3; Length 28;

Matches 16; Conservative 4; Mismatches 2; Indels 7; Gaps 1;

Qy 1 KOAEDKVKASREAKQVEKALEQLEDKVK 29  
Db 4 KOAEDKV-----KQLEDKVBELEQDKVK 25

## RESULT 14

US-08-817-811-52  
Sequence 52, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Relif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-52

Query Match 48.9%; Score 67; DB 3; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.15;  
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKQVEKALEQLEDKVK 29  
Db 1 KQLEDKVMAQDTADRLTEKLNQLEDKVK 29

## RESULT 15

US-08-817-811-15  
Sequence 15, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Relif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-15

Query Match 47.8%; Score 65.5; DB 3; Length 28;  
Best Local Similarity 58.6%; Pred. No. 0.21;  
Matches 17; Conservative 2; Mismatches 3; Indels 7; Gaps 1;

Qy 1 KOAEDKVKASREAKQVEKALEQLEDKVK 29  
Db 5 KOAERDASREAK-----QLEQDKVK 26

## RESULT 16

US-08-817-811-74  
Sequence 74, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Relif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996

```

/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Highlander, Steven L.
/ REGISTRATION NUMBER: 37,642
/ REFERENCE/DOCKET NUMBER: FBRC:005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512/418-3000
/ TELEFAX: 512/474-7577
/ INFORMATION FOR SEQ ID NO: 74:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 29 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/
US-08-817-811-74

Query Match 47.4%; Score 65; DB 3; Length 29;
Best Local Similarity 48.3%; Pred. No. 0.25;
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KQAEKVKASREAKKQVEKALEQLEDKVK 29
Db 1 KQLEKVKQAEKXKLNIGRKQLKQLODKVK 29

RESULT 17
US-08-817-811-1
/ Sequence 1, Application US/08817811
/ Patent No. 6174528
/ GENERAL INFORMATION:
/ APPLICANT: Cooper, Juan A.
/ APPLICANT: Reif, Wendy A.
/ APPLICANT: Good, Michael F.
/ APPLICANT: Saul, Allan J.
/ TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
/ NUMBER OF SEQUENCES: 97
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Arnold, White & Durkee
/ STREET: P.O. Box 4433
/ CITY: Houston
/ STATE: Texas
/ COUNTRY: USA
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/817, 811
/ FILING DATE: 14-APR-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO 96/11944
/ FILING DATE: 25-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Highlander, Steven L.
/ REGISTRATION NUMBER: 37,642
/ REFERENCE/DOCKET NUMBER: FBRC:005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512/418-3000
/ TELEFAX: 512/474-7577
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/
US-08-817-811-1

Query Match 46.7%; Score 64; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 9 ASREAKKQVEKALE 22
Db 7 ASREAKKQVEKALE 20

RESULT 18
US-08-817-811-73
/ Sequence 73, Application US/08817811
/ Patent No. 6174528
/ GENERAL INFORMATION:
/ APPLICANT: Cooper, Juan A.
/ APPLICANT: Reif, Wendy A.
/ APPLICANT: Good, Michael F.
/ APPLICANT: Saul, Allan J.
/ TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES
/ NUMBER OF SEQUENCES: 97
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Arnold, White & Durkee
/ STREET: P.O. Box 4433
/ CITY: Houston
/ STATE: Texas
/ COUNTRY: USA
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/817, 811
/ FILING DATE: 14-APR-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO 96/11944
/ FILING DATE: 25-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Highlander, Steven L.
/ REGISTRATION NUMBER: 37,642
/ REFERENCE/DOCKET NUMBER: FBRC:005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512/418-3000
/ TELEFAX: 512/474-7577
/ INFORMATION FOR SEQ ID NO: 73:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 29 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/
US-08-817-811-73

Query Match 46.7%; Score 64; DB 3; Length 29;
Best Local Similarity 48.3%; Pred. No. 0.33;
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KQAEKVKASREAKKQVEKALEQLEDKVK 29
Db 1 KQLEKVKQAEKXKLNIGRKQLKQLODKVK 29

RESULT 19
US-08-182-175A-87
/ Sequence 87, Application US/08182175A
/ Patent No. 5559223
/ GENERAL INFORMATION:
/ APPLICANT: Saverio Carl Falco
/ APPLICANT: Sharon J. Keeler
/ APPLICANT: Janet A. Rice
/ TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing P
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E.I. du Pont de Nemours and Company
```

STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,175A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamechey Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-182-175A-87

Query Match 46.7%; Score 64; DB 1; Length 72;  
Best Local Similarity 41.4%; Pred. No. 0.84;  
Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KOAEDKYKASREAKQVEKALEQLEDKYK 29  
Db 6 KKLSEKMKAMEBKMKWLEBKMKKLEBKXK 34

RESULT 20  
PCT-US92-06412-87  
Sequence 87, Application PC/TUS9206412  
GENERAL INFORMATION:  
APPLICANT: Severio Carl Falco  
APPLICANT: Sharon J. Keeler  
APPLICANT: Janet A. Rice  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing F  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E.I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06412  
FILING DATE: 19920807  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamechey Floyd  
REGISTRATION NUMBER: 33,692

REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-06412-87

Query Match 46.7%; Score 64; DB 5; Length 72;  
Best Local Similarity 41.4%; Pred. No. 0.84;  
Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KOAEDKYKASREAKQVEKALEQLEDKYK 29  
Db 6 KKLSEKMKAMEBKMKWLEBKMKKLEBKXK 34

RESULT 21  
US-09-543-681A-5390  
Sequence 5390, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETTON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709,1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
PRIOR FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 5390  
LENGTH: 361  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-5390

Query Match 46.7%; Score 64; DB 4; Length 361;  
Best Local Similarity 48.3%; Pred. No. 4.4;  
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 KOAEDKYKASREAKQVEKALEQLEDKYK 29  
Db 165 EQADAKAKAEAKKQAEAKQAEAK 193

RESULT 22  
US-08-817-811-79  
Sequence 79, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Relf, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-79

Query Match 46.0%; Score 63; DB 3; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.43;  
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 KQADKYKASREAKQVEKALBQLEDKVK 29  
Db 1 KQLEKVKQAEKLNQKQLAQLODKVK 29

## RESULT 23

US-08-182-175A-105  
Sequence 105, Application US/08182175A  
Patent No. 5559223  
GENERAL INFORMATION:  
APPLICANT: Saverio Carl Falco  
APPLICANT: Sharon J. Keeler  
APPLICANT: Janet A. Rice  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing F  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,175A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamethy Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-182-175A-105

Query Match 46.0%; Score 63; DB 1; Length 107;  
Best Local Similarity 37.9%; Pred. No. 1.6;  
Matches 11; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KQADKYKASREAKQVEKALBQLEDKVK 29  
Db 27 KQLEKVKQAEKLNQKQLAQLODKVK 55

## RESULT 24

US-08-474-633A-92  
Sequence 92, Application US/08474633A  
Patent No. 5773691  
GENERAL INFORMATION:  
APPLICANT: E. I. du Pont de Nemours and  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
TITLE OF INVENTION: OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. du Pont de Nemours  
ADDRESS: AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,633A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIGGEL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 92:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-474-633A-92

Query Match 46.0%; Score 63; DB 1; Length 107;  
Best Local Similarity 37.9%; Pred. No. 1.6;  
Matches 11; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KQADKYKASREAKQVEKALBQLEDKVK 29  
Db 27 KQLEKVKQAEKLNQKQLAQLODKVK 55

## RESULT 25

US-08-823-771-92  
Sequence 92, Application US/08823771  
Patent No. 6459019



SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-13

Query Match 45.3%; Score 62; DB 3; Length 28;  
Best Local Similarity 51.4%; Pred. No. 0.54;  
Matches 18; Conservative 2; Mismatches 1; Indels 14; Gaps 2;

Qy 2 QAEDEVK-----ASREAKKQVEKALEQLEDEKVK 29  
Db 1 QLEDEVKQLRRDLASREAK-----BELQDEKVK 28

RESULT 28  
US-08-817-811-78

Sequence 78, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817.811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-78

Query Match 45.3%; Score 62; DB 3; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.56;  
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 KOAEDKVASREAKKQVEKALEQLEDEKVK 29  
Db 1 KOLEBKVKQATEKLNIGKQQLAQLEQDKVK 29

RESULT 29  
US-08-182-175A-57

Sequence 57, Application US/08182175A

Patent No. 5559223  
GENERAL INFORMATION:  
APPLICANT: Saverio Carl Falco  
APPLICANT: Sharon J. Keeler  
APPLICANT: Janet A. Rice  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing P  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E.I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182.175A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743.006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Kamechly Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420

INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-182-175A-57

Query Match 45.3%; Score 62; DB 1; Length 77;  
Best Local Similarity 41.4%; Pred. No. 1.5;  
Matches 12; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KOAEDKVASREAKKQVEKALEQLEDEKVK 29  
Db 27 KAMEBKVKAMEKVKAMEBKVKAMEBKVK 55

RESULT 30  
US-08-474-633A-75

Sequence 75, Application US/08474633A  
Patent No. 5773591  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
TITLE OF INVENTION: METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
TITLE OF INVENTION: OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898

US-08-474-633A-75

Sequence 75, Application US/08474633A  
Patent No. 5773591  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
TITLE OF INVENTION: METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
TITLE OF INVENTION: OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898

US-08-474-633A-75

TELECOMMUNICATION INFORMATION  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164

QY 1 KQAEADVKAASREAKKQVEKALEQLEDDKVK 29  
Db 27 KAMEEKTLKAMEEKLKAMEEKLKQKMEELK 55

### RESULT 33

US-08-182-175A-49  
Sequence 49, Application US/08182175A  
Patent No. 5559223  
GENERAL INFORMATION:  
APPLICANT: Saverio Carl Falco  
APPLICANT: Sharon J. Keeler  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing F  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E.I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,175A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamechy Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-182-175A-49  
Query Match 44.5%; Score 61; DB 1; Length 28;  
Best Local Similarity 42.3%; Pred. No. 0.7;  
Matches 11; Conservative 10; Mismatches 5; Indels 0; Gaps 0;  
Qy 4 EDKVASREAKQVEKALEQLEDKVK 29  
Db 2 EKKKAMEKTKKMEKTKKMEKMK 27  
RESULT 34  
US-08-474-633A-58  
Sequence 58, Application US/08474633A  
Patent No. 5773691  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
TITLE OF INVENTION: OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.

ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,633A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIEGEL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-474-633A-58  
Query Match 44.5%; Score 61; DB 1; Length 28;  
Best Local Similarity 42.3%; Pred. No. 0.7;  
Matches 11; Conservative 10; Mismatches 5; Indels 0; Gaps 0;  
Qy 4 EDKVASREAKQVEKALEQLEDKVK 29  
Db 2 EKKKAMEKTKKMEKTKKMEKMK 27  
RESULT 35  
US-08-823-771-58  
Sequence 58, Application US/08823771  
Patent No. 6459019  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/823,771  
FILING DATE: 24-Mar-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/474,633  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIEGEL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931

TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 58:  
US-08-823-771-58

Query Match 44.5%; Score 61; DB 4; Length 28;  
Best Local Similarity 42.3%; Pred. No. 0.7;  
Matches 11; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 4 EDKVKASREAKKQVEKALEQLEDKVK 29  
DB 2 EEKMKAMEKIKKMEKELKKMEKKK 27

RESULT 36  
PCT-US92-06412-49  
Sequence 49, Application PC/TUS9206412  
GENERAL INFORMATION:  
APPLICANT: Saverio Carl Falco  
APPLICANT: Sharon J. Keeler  
APPLICANT: Janet A. Rice  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing F  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06412  
FILING DATE: 19920807  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamechy Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-06412-49

Query Match 44.5%; Score 61; DB 5; Length 28;  
Best Local Similarity 42.3%; Pred. No. 0.7;  
Matches 11; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 4 EDKVKASREAKKQVEKALEQLEDKVK 29  
DB 2 EEKMKAMEKIKKMEKELKKMEKKK 27

RESULT 37  
US-08-817-811-71  
Sequence 71, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: PIRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 71:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-71

Query Match 44.5%; Score 61; DB 3; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.73;  
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 29  
DB 1 KQLEKVKARLTKELKNIQKQKQLEQDKVK 29

RESULT 38  
US-08-817-811-72  
Sequence 72, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA

ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-72

Query Match 44.5%; Score 61; DB 3; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.73;  
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 KOAEKVKASREAKKQVEKALEQLEDKVK 29  
DB 1 KQLEEKVKQTEKKNIQKQKQLOQDKVK 29

RESULT 39  
US-08-817-811-75  
Sequence 75, Application US/08817811  
Patent No. 6174528  
GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-75

Query Match 44.5%; Score 61; DB 3; Length 29;  
Best Local Similarity 44.8%; Pred. No. 0.73;  
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 KOAEKVKASREAKKQVEKALEQLEDKVK 29  
DB 1 KQLEEKVKQTEKKNIQKQKQLOQDKVK 29

RESULT 40

US-08-817-811-89  
Sequence 89, Application US/08817811  
Patent No. 6174528

GENERAL INFORMATION:  
APPLICANT: Cooper, Juan A.  
APPLICANT: Reif, Wendy A.  
APPLICANT: Good, Michael F.  
APPLICANT: Saul, Allan J.  
TITLE OF INVENTION: SYNTHETIC PEPTIDES AND VACCINES  
TITLE OF INVENTION: COMPRISING SAME  
NUMBER OF SEQUENCES: 97  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,811  
FILING DATE: 14-APR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 96/11944  
FILING DATE: 25-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: FBRC:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-817-811-89

Query Match 44.5%; Score 61; DB 3; Length 29;  
Best Local Similarity 48.3%; Pred. No. 0.73;  
Matches 14; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 KOAEKVKASREAKKQVEKALEQLEDKVK 29  
DB 1 KQLEEKVKQTEKKNIQKQKQLOQDKVK 29

Search completed: June 13, 2005, 20:58:37  
Job time : 41.4651 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 20:52:55 ; Search time 141.628 Seconds  
(without alignments)  
78.492 Million cell updates/sec

Title: US-10-706-275-2

Perfect score: 137  
Sequence: 1 KQAEKVKASREAKKVEKALEQEDKVK 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PC105\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	137	100.0	29 17 US-10-706-275-2	Sequence 2, Appli
2	137	100.0	29 17 US-10-706-275-15	Sequence 15, Appl
3	125	91.2	28 17 US-10-706-275-12	Sequence 12, Appl
4	119	86.9	28 17 US-10-706-275-13	Sequence 13, Appl
5	112	81.8	28 17 US-10-706-275-14	Sequence 14, Appl
6	110	80.3	28 17 US-10-706-275-11	Sequence 11, Appl
7	96	70.1	28 17 US-10-706-275-10	Sequence 10, Appl
8	81	59.1	28 17 US-10-706-275-9	Sequence 9, Appl
9	72	52.6	443 8 US-08-325-278-6	Sequence 6, Appli
10	72	52.6	553 16 US-10-474-792-672	Sequence 672, App
11	72	52.6	558 17 US-10-732-923-3295	Sequence 3295, Ap

12	71	51.8	254 13 US-10-141-627-4	Sequence 4, Appli
13	71	51.8	284 13 US-10-141-627-6	Sequence 6, Appli
14	67	48.9	389 15 US-10-282-122A-67145	Sequence 67145, A
15	65.5	47.8	28 17 US-10-706-275-8	Sequence 8, Appli
16	64	46.7	14 17 US-10-706-275-1	Sequence 1, Appli
17	64	46.7	20 13 US-10-044-034-32	Sequence 22, Appl
18	64	46.7	20 17 US-10-706-275-5	Sequence 5, Appli
19	63	46.0	107 14 US-10-023-066A-92	Sequence 92, Appl
20	62	45.3	107 17 US-10-804-678-92	Sequence 92, Appl
21	62	45.3	28 17 US-10-706-275-6	Sequence 6, Appli
22	62	45.3	77 14 US-10-023-066A-75	Sequence 75, Appl
23	62	45.3	77 17 US-10-804-678-75	Sequence 75, Appl
24	61	44.5	28 14 US-10-023-066A-58	Sequence 58, Appl
25	61	44.5	28 17 US-10-804-678-58	Sequence 58, Appl
26	61	44.5	37 14 US-10-023-066A-85	Sequence 85, Appl
27	61	44.5	37 14 US-10-023-066A-86	Sequence 86, Appl
28	61	44.5	37 17 US-10-804-678-85	Sequence 85, Appl
29	61	44.5	37 17 US-10-804-678-86	Sequence 86, Appl
30	61	44.5	56 14 US-10-023-066A-77	Sequence 77, Appl
31	61	44.5	56 17 US-10-804-678-77	Sequence 77, Appl
32	61	44.5	145 16 US-10-437-963-146357	Sequence 146357,
33	61	44.5	145 16 US-10-437-963-146368	Sequence 146368,
34	60	43.8	35 14 US-10-023-066A-62	Sequence 62, Appl
35	60	43.8	35 17 US-10-804-678-62	Sequence 62, Appl
36	60	43.8	42 14 US-10-023-066A-34	Sequence 34, Appl
37	60	43.8	42 17 US-10-804-678-34	Sequence 34, Appl
38	60	43.8	49 14 US-10-023-066A-30	Sequence 30, Appl
39	60	43.8	49 14 US-10-023-066A-32	Sequence 32, Appl
40	60	43.8	49 14 US-10-023-066A-54	Sequence 54, Appl
41	60	43.8	49 17 US-10-804-678-30	Sequence 30, Appl
42	60	43.8	49 17 US-10-804-678-32	Sequence 32, Appl
43	60	43.8	49 17 US-10-804-678-54	Sequence 54, Appl
44	60	43.8	56 14 US-10-023-066A-79	Sequence 79, Appl
45	60	43.8	56 17 US-10-804-678-79	Sequence 79, Appl

## ALIGNMENTS

RESULT 1  
US-10-706-275-2  
; Sequence 2, Application US/10706275  
; Publication No. US2005002956A1  
; GENERAL INFORMATION:  
; APPLICANT: ID Biomedical Corporation of Quebec  
; APPLICANT: The Council of the Queensland Institute of Medical Research  
; APPLICANT: Lowell, George H.  
; APPLICANT: Burt, David S.  
; APPLICANT: White, Gregory L.  
; APPLICANT: Good, Michael F.  
; APPLICANT: Batzloff, Michael R.  
; APPLICANT: leanderson, Tomas B.  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: 021989-000710US  
; CURRENT APPLICATION NUMBER: US/10/706, 275  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: US 60/426, 409  
; PRIOR FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: AU 2002302132  
; PRIOR FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: antigenic peptide derivative of S. pyogenes with flanking sequenc  
; OTHER INFORMATION: es  
US-10-706-275-2  
Query Match 100.0%; Score 137, DB 17, Length 29,  
Best Local Similarity 100.0%; Pred. No. 7,6e-09;

```
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 29
    |||||
Db 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 29

RESULT 2
US-10-706-275-15
; Sequence 15, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Good, Michael F.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: AU 2002302132
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antigenic peptide sequence derivative of p145 baase
US-10-706-275-15

Query Match 100.0%; Score 137; DB 17; Length 29;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 29
    |||||
Db 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 29

RESULT 3
US-10-706-275-12
; Sequence 12, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: AU 2002302132
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 28
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antigenic peptide sequence derivative of p145 baase
US-10-706-275-12

Query Match 91.2%; Score 125; DB 17; Length 28;
Best Local Similarity 92.9%; Pred. No. 1.7e-07;
Matches 26; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 28
    |||||
Db 1 KOAEDKVKASREAKKQVEKALEQLEDKVK 28

RESULT 4
US-10-706-275-13
; Sequence 13, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Good, Michael F.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: AU 2002302132
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antigenic peptide sequence derivative of p145 baase
US-10-706-275-13

Query Match 86.9%; Score 119; DB 17; Length 28;
Best Local Similarity 89.3%; Pred. No. 8.1e-07;
Matches 25; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QAEDKVKASREAKKQVEKALEQLEDKVK 29
    |||||
Db 1 QAEDKVKASREAKKQVEKALEQLEDKVK 28

RESULT 5
US-10-706-275-14
; Sequence 14, Application US/10706275
; Publication No. US20050002956A1
; GENERAL INFORMATION:
; APPLICANT: ID Biomedical Corporation of Quebec
; APPLICANT: The Council of the Queensland Institute of Medical Research
; APPLICANT: Lowell, George H.
; APPLICANT: Burt, David S.
; APPLICANT: White, Gregory L.
; APPLICANT: Batzloff, Michael R.
; APPLICANT: Leanderson, Tomas B.
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 021989-000710US
; CURRENT APPLICATION NUMBER: US/10/706,275
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,409
; PRIOR FILING DATE: 2002-11-15
```

PRIOR APPLICATION NUMBER: AU 2002302132  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 14  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-14

Query Match 81.8%; Score 112; DB 17; Length 28;  
Best Local Similarity 88.9%; Pred. No. 5.1e-06;  
Matches 24; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 AEDKVKASREAKKQVEKALEQLEDEK 29  
DB 1 AEDKVKQLEAKKQVEKALEQLEDEK 27

RESULT 6  
US-10-706-275-11  
Sequence 11, Application US/10706275  
Publication No. US2005002956A1  
GENERAL INFORMATION:  
APPLICANT: ID Biomedical Corporation of Quebec  
APPLICANT: The Council of the Queensland Institute of Medical Research  
APPLICANT: Lowell, George H.  
APPLICANT: Burt, David S.  
APPLICANT: White, Gregory L.  
APPLICANT: Batzloff, Michael F.  
APPLICANT: Leanderson, Tomas B.  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 021989-000710US  
CURRENT APPLICATION NUMBER: US/10/706,275  
CURRENT FILING DATE: 2003-11-13  
PRIOR APPLICATION NUMBER: US 60/426,409  
PRIOR FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: AU 2002302132  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 11  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-11

Query Match 80.3%; Score 110; DB 17; Length 28;  
Best Local Similarity 85.2%; Pred. No. 8.5e-06;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KOAEDKVKASREAKKQVEKALEQLEDEK 27  
DB 2 KOAEDKVDASREAKKQVEKVKQLEDEK 28

RESULT 7  
US-10-706-275-10  
Sequence 10, Application US/10706275  
Publication No. US2005002956A1  
GENERAL INFORMATION:  
APPLICANT: ID Biomedical Corporation of Quebec  
APPLICANT: The Council of the Queensland Institute of Medical Research  
APPLICANT: Lowell, George H.  
APPLICANT: Burt, David S.  
APPLICANT: White, Gregory L.  
APPLICANT: Good, Michael F.  
APPLICANT: Batzloff, Michael R.

APPLICANT: Leanderson, Tomas B.  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 021989-000710US  
CURRENT APPLICATION NUMBER: US/10/706,275  
CURRENT FILING DATE: 2003-11-13  
PRIOR APPLICATION NUMBER: US 60/426,409  
PRIOR FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: AU 2002302132  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 10  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-10

Query Match 70.1%; Score 96; DB 17; Length 28;  
Best Local Similarity 76.9%; Pred. No. 0.00033;  
Matches 20; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 KOAEDKVKASREAKKQVEKALEQLEDEK 26  
DB 3 KOAEDKLDASREAKKQVEKVKQLEDEK 28

RESULT 8  
US-10-706-275-9  
Sequence 9, Application US/10706275  
Publication No. US2005002956A1  
GENERAL INFORMATION:  
APPLICANT: ID Biomedical Corporation of Quebec  
APPLICANT: The Council of the Queensland Institute of Medical Research  
APPLICANT: Lowell, George H.  
APPLICANT: Burt, David S.  
APPLICANT: White, Gregory L.  
APPLICANT: Batzloff, Michael F.  
APPLICANT: Leanderson, Tomas B.  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 021989-000710US  
CURRENT APPLICATION NUMBER: US/10/706,275  
CURRENT FILING DATE: 2003-11-13  
PRIOR APPLICATION NUMBER: US 60/426,409  
PRIOR FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: AU 2002302132  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-9

Query Match 59.1%; Score 81; DB 17; Length 28;  
Best Local Similarity 68.0%; Pred. No. 0.017;  
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 KOAEDKVKASREAKKQVEKALEQLEDEK 25  
DB 4 KOAEDDLASREAKKQVODKVKQLEDEK 28

RESULT 9  
US-08-325-278-6  
Sequence 6, Application US/08325278  
Publication No. US20030027283A1  
GENERAL INFORMATION:

```

; APPLICANT: Bjvreck, Lars
; APPLICANT: Strybing, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-325-278-6
;
Query Match          52.6%; Score 72; DB 8; Length 443;
Best Local Similarity 45.2%; Pred. No. 3.3;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQVEKALEQLEBDKV 28
Db 272 DKVKEKQISDASRQRLRDLDASREAKQVEKALEANSKL 313

RESULT 10
US-10-474-792-672
; Sequence 672, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Louie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 672
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
;
US-10-474-792-672
;
Query Match          52.6%; Score 72; DB 16; Length 553;
Best Local Similarity 45.2%; Pred. No. 4.2;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASRAKQVEKALEQLEBDKV 28
Db 382 DKVKEKQISDASRQRLRDLDASREAKQVEKALEANSKL 423
```

```

RESULT 11
US-10-732-923-3295
; Sequence 3295, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3295
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
;
US-10-732-923-3295
;
Query Match          52.6%; Score 72; DB 17; Length 558;
Best Local Similarity 50.0%; Pred. No. 4.2;
Matches 18; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 1 KQAE-----KVKSREAKQVEKALEQLEBDKV 28
Db 393 KQTSASRQGLRDLDSREAKQVEKALEANSKL 428

RESULT 12
US-10-141-627-4
; Sequence 4, Application US/10141627
; Publication No. US20020176863A1
; GENERAL INFORMATION:
; APPLICANT: Dale, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
; FILE REFERENCE: 48112.404C3
; CURRENT APPLICATION NUMBER: US/10/141,627
; CURRENT FILING DATE: 2002-05-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: An antigen of M5 and a carrier of the
; OTHER INFORMATION: COOH-terminal portion of M5
;
US-10-141-627-4
;
Query Match          51.8%; Score 71; DB 13; Length 254;
Best Local Similarity 75.0%; Pred. No. 2.4;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 ASREAKQVEKALEQLEBDKV 28
Db 105 ASREAKQVEKALEANSKL 124

RESULT 13
US-10-141-627-6
; Sequence 6, Application US/10141627
; Publication No. US20020176863A1
; GENERAL INFORMATION:
; APPLICANT: Dale, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
; FILE REFERENCE: 48112.404C3
; CURRENT APPLICATION NUMBER: US/10/141,627
; CURRENT FILING DATE: 2002-05-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
;
US-10-141-627-6
```

SEQ ID NO 6  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: An antigen of three fragments of M5 and a carrier  
US-10-141-627-6

Query Match 51.8%; Score 71; DB 13; Length 284;  
Best Local Similarity 75.0%; Pred. No. 2.7;  
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ASREAKQVERKLEQLQEDKVK 28  
Db 135 ASREAKQVERKLEANSKL 154

RESULT 14  
US-10-282-122A-67145  
Sequence 67145, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haseelbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyakind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Cart, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT FILING DATE: 2003-02-20  
PRIOR FILING DATE: 2003-02-20  
PRIOR FILING DATE: 2000-03-21  
PRIOR FILING DATE: 2000-03-21  
PRIOR FILING DATE: 2000-05-23  
PRIOR FILING DATE: 2000-05-26  
PRIOR FILING DATE: 2000-07-27  
PRIOR FILING DATE: 2000-09-06  
PRIOR FILING DATE: 2000-09-09  
PRIOR FILING DATE: 2000-09-09  
PRIOR FILING DATE: 2000-10-23  
PRIOR FILING DATE: 2000-10-23  
PRIOR FILING DATE: 2000-11-27  
PRIOR FILING DATE: 2000-12-22  
PRIOR FILING DATE: 2001-02-09  
PRIOR FILING DATE: 2001-02-09  
PRIOR FILING DATE: 2001-02-15  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 67145  
LENGTH: 389  
TYPE: PRT  
ORGANISM: Pasteurella multocida  
US-10-282-122A-67145

Query Match 48.9%; Score 67; DB 15; Length 389;  
Best Local Similarity 51.6%; Pred. No. 11;  
Matches 16; Conservative 5; Mismatches 8; Indels 2; Gaps 1;  
Qy 1 KQAEED--KVKASREAKQVERKLEQLQEDKVK 29

Db 175 KQAESEAKAKAAEAKRKAKAKAAEAKAK 205

RESULT 15  
US-10-706-275-8  
Sequence 8, Application US/10706275  
Publication No. US2005002956A1  
GENERAL INFORMATION:  
APPLICANT: ID Biomedical Corporation of Quebec  
APPLICANT: The Council of the Queensland Institute of Medical Research  
APPLICANT: Lowell, George H.  
APPLICANT: Burt, David S.  
APPLICANT: White, Gregory L.  
APPLICANT: Good, Michael F.  
APPLICANT: Batzloff, Michael R.  
APPLICANT: Leanderson, Tomas B.  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 021989-000710US  
CURRENT FILING DATE: 2003-11-13  
PRIOR FILING DATE: 2003-11-13  
PRIOR FILING DATE: 2002-11-15  
PRIOR FILING DATE: 2002-11-15  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of p145 base  
US-10-706-275-8

Query Match 47.8%; Score 65.5; DB 17; Length 28;  
Best Local Similarity 58.6%; Pred. No. 0.97;  
Matches 17; Conservative 2; Mismatches 3; Indels 7; Gaps 1;

Qy 1 KQAEKVKASREAKQVERKLEQLQEDKVK 29  
Db 5 KQAEKVDASREAK-----QLQEDKVK 26

RESULT 16  
US-10-706-275-1  
Sequence 1, Application US/10706275  
Publication No. US2005002956A1  
GENERAL INFORMATION:  
APPLICANT: ID Biomedical Corporation of Quebec  
APPLICANT: The Council of the Queensland Institute of Medical Research  
APPLICANT: Lowell, George H.  
APPLICANT: Burt, David S.  
APPLICANT: White, Gregory L.  
APPLICANT: Good, Michael F.  
APPLICANT: Batzloff, Michael R.  
APPLICANT: Leanderson, Tomas B.  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 021989-000710US  
CURRENT FILING DATE: 2003-11-13  
PRIOR FILING DATE: 2003-11-13  
PRIOR FILING DATE: 2002-11-15  
PRIOR FILING DATE: 2002-11-15  
PRIOR FILING DATE: 2002-11-15  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: antigenic peptide sequence derivative of S. pyogenes

US-10-706-275-1

Query Match 46.7%; Score 64; DB 17; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.69; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALE 22

Db 1 ASREAKKQVEKALE 14

RESULT 17

US-10-044-034-22  
Sequence 22, Application US/10044034  
Publication No. US20020169264A1

GENERAL INFORMATION:

APPLICANT: JACKSON, DAVID C.

APPLICANT: O'BRIEN-STIMPSON, NEIL M.

APPLICANT: BROWN, LORENA E.

APPLICANT: EDE, NICHOLAS J.

APPLICANT: BRANDT, EVELYN R.

APPLICANT: GOOD, MICHAEL F.

TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES

FILE REFERENCE: FERC:006

CURRENT APPLICATION NUMBER: US/10/044.034

CURRENT FILING DATE: 2002-01-11

PRIOR APPLICATION NUMBER: P05071

PRIOR FILING DATE: 1997-02-11

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 22

LENGTH: 20

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Peptides

US-10-044-034-22

Query Match 46.7%; Score 64; DB 13; Length 20;

Best Local Similarity 100.0%; Pred. No. 1; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALE 22

Db 7 ASREAKKQVEKALE 20

RESULT 18

US-10-706-275-5  
Sequence 5, Application US/10706275  
Publication No. US2005002956A1

GENERAL INFORMATION:

APPLICANT: ID Biomedical Corporation of Quebec

APPLICANT: The Council of the Queensland Institute of Medical Research

APPLICANT: Lowell, George H.

APPLICANT: Burt, David S.

APPLICANT: White, Gregory L.

APPLICANT: Good, Michael F.

APPLICANT: Batzloff, Thomas R.

APPLICANT: Leanderson, Tomas B.

TITLE OF INVENTION: Vaccine

FILE REFERENCE: 021989-000710US

CURRENT APPLICATION NUMBER: US/10/706,275

CURRENT FILING DATE: 2003-11-13

PRIOR APPLICATION NUMBER: US 60/426,409

PRIOR FILING DATE: 2002-11-15

PRIOR APPLICATION NUMBER: AU 2002302132

PRIOR FILING DATE: 2002-11-15

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 20

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE: antigenic peptide sequence p145

OTHER INFORMATION:

US-10-706-275-5

Query Match 46.7%; Score 64; DB 17; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ASREAKKQVEKALE 22

Db 7 ASREAKKQVEKALE 20

RESULT 19

US-10-023-066A-92  
Sequence 92, Application US/10023066A  
Publication No. US20030056242A1

GENERAL INFORMATION:

APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY

TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR

INCREASING THE LYSINE AND

THREONINE CONTENT OF THE SEEDS OF

PLANTS

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS

AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD VERSION 2.0C

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/023,066A

FILING DATE: 29-Apr-2002

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: BARBARA C. SIEGEL

REGISTRATION NUMBER: 30,684

REFERENCE/DOCKET NUMBER: BB-1037-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-4931

TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 92:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 92:

US-10-023-066A-92

Query Match 46.0%; Score 63; DB 14; Length 107;  
Best Local Similarity 37.9%; Pred. No. 7.7;  
Matches 11; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 KQADKYKASREAKKQVEKALEQLEDKVK 29

Db 27 KKLSEKKQVMEKKKKLSEKKKAKMEDKVK 55

RESULT 20

US-10-804-678-92  
Sequence 92, Application US/10804678

```

1 Publication No. US2005000530A1
2 GENERAL INFORMATION:
3 APPLICANT: EPELANUM, SABINE URSULA
4 FALCO, SAVERIO CARL
5 MCDEVITT, RAYMOND ERVIN, III
6 TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR
7 INCREASING THE LYSINE CONTENT OF
8 THE SEEDS OF PLANTS
9 NUMBER OF SEQUENCES: 132
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
12 STREET: 1007 MARKET STREET
13 CITY: WILMINGTON
14 STATE: DELAWARE
15 COUNTRY: U.S.A.
16 ZIP: 19898
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: DISKETTE, 3.50 INCH
19 COMPUTER: IBM PC COMPATIBLE
20 OPERATING SYSTEM: MICROSOFT OFFICE 97
21 SOFTWARE: MICROSOFT WINDOWS 95
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/10/804,678
24 FILING DATE: 19-Mar-2004
25 CLASSIFICATION: <Unknown>
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US/09/049,304
28 FILING DATE: 27-Mar-1998
29 APPLICATION NUMBER: 08/824,627
30 FILING DATE: MARCH 27, 1997
31 ATTORNEY/AGENT INFORMATION:
32 NAME: CHRISTENBURY, LYNN M.
33 REGISTRATION NUMBER: 30,971
34 REFERENCE/DOCKET NUMBER: BB-1037-F
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 302-992-5481
37 TELEFAX: 302-892-7949
38 TELEX: 835420
39 INFORMATION FOR SEQ ID NO: 92:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 107 amino acids
42 TYPE: amino acid
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 SEQUENCE DESCRIPTION: SEQ ID NO: 92:
46 US-10-804-678-92
47
48 Query Match 46.0%; Score 63; DB 17; Length 107;
49 Best Local Similarity 37.9%; Pred.No. 7.7;
50 Matches 11; Conservative 10; Mismatches 8; Indels 0; Gaps 0.
51
52 Oy 1 KOAEKVKASREAKKQVEKALEQLEDKYR 29
53 |::|::||::|::|::|::|
54 Db 27 KKLSEKKMKMEKKKKLSEKKMAEKDKYK 55
55
56 RESULT 21
57 US-10-706-275-6
58 Sequence 6, Application US/10706275
59 Publication No. US20050002936A1
60 GENERAL INFORMATION:
61 APPLICANT: ID Biomedical Corporation of Quebec
62 APPLICANT: The Council of the Queensland Institute of Medical Research
63 APPLICANT: Lowell, George H.
64 APPLICANT: Burt, David S.
65 APPLICANT: White, Gregory L.
66 APPLICANT: Good, Michael F.
67 APPLICANT: Batzloff, Michael R.
68 APPLICANT: Leanderson, Tomas B.
69 TITLE OF INVENTION: Vaccine
70 FILE REFERENCE: 021989-000710US
71 CURRENT APPLICATION NUMBER: US/10/706,275
72 CURRENT FILING DATE: 2003-11-13

```

```

: PRIOR APPLICATION NUMBER: US 60/426,409
: PRIOR FILING DATE: 2002-11-15
: PRIOR APPLICATION NUMBER: AU 2002302132
: PRIOR FILING DATE: 2002-11-15
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 6
: LENGTH: 28
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: antigenic peptide sequence derivative of p145 base
US-10-706-275-6

Query Match      45.3%; Score 62; DB 17; Length 28;
Best Local Similarity 51.4%; Pred. No. 2.4;
Matches 18; Conservative 2; Mismatches 1; Indels 14; Gaps 2

Cy      2 QAEKVK-----ASREAKQVEKALEQLDVK 29
      1 QLEDVKQLRRDLDSREAK-----EELQDKVK 28

RESULT 22
US-10-023-066A-75
: Sequence 75, Application US/10023066A
: Publication No. US20030056242A1
: GENERAL INFORMATION:
: APPLICANT: E. I. DU PONT DE NEMOURS AND
: COMPANY
: TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR
: INCREASING THE LYSINE AND
: THREONINE CONTENT OF THE SEEDS OF
: PLANTS
: NUMBER OF SEQUENCES: 107
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: E. I. DU PONT DE NEMOURS
: AND COMPANY
: STREET: 1007 MARKET STREET
: CITY: WILMINGTON
: STATE: DELAWARE
: COUNTRY: U.S.A.
: ZIP: 19898
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY DISK
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: MICROSOFT WORD VERSION 2.0C
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/023,066A
: FILING DATE: 29-Apr-2002
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: BARBARA C. SIEGEL
: REGISTRATION NUMBER: 30,684
: REFERENCE/DOCKET NUMBER: BB-1037-C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 302-992-4931
: TELEFAX: 302-773-0164
: TELEX: 835420
: INFORMATION FOR SEQ ID NO: 75:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 77 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-10-023-066A-75

Query Match      45.3%; Score 62; DB 14; Length 77;
Best Local Similarity 41.4%; Pred. No. 7.1;
Matches 12; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```

Qy 1 KOAEDKVKASREAKQVEKALEQLEDDKVK 29  
Db 27 KAMEBKVKAMEBKVKAMEBKVKAMEBKVK 55

## RESULT 23

US-10-804-678-75  
Sequence 75, Application US/10804678  
Publication No. US20050005330A1  
GENERAL INFORMATION:

APPLICANT: EPELBAUM, SABINE URSULA

FALCO, SAVERIO CARL

MCDEVITT, RAYMOND ERVIN, III

TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR

INCREASING THE LYSINE CONTENT OF

THE SEEDS OF PLANTS

NUMBER OF SEQUENCES: 132

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: MICROSOFT OFFICE 97

SOFTWARE: MICROSOFT WINDOWS 95

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/804,678

FILING DATE: 19-Mar-2004

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/049,304

FILING DATE: 27-Mar-1998

APPLICATION NUMBER: 08/824,627

FILING DATE: MARCH 27, 1997

ATTORNEY/AGENT INFORMATION:

NAME: CHRISTENBURY, LYNN M.

REGISTRATION NUMBER: 30,971

REFERENCE/DOCKET NUMBER: BB-1037-F

TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-5481

TELEFAX: 302-892-7949

TELEX: 835420

INFORMATION FOR SEQ ID NO: 75:

SEQUENCE CHARACTERISTICS:

LENGTH: 77 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 75:

US-10-804-678-75

Query Match 45.3%; Score 62; DB 17; Length 77;  
Best Local Similarity 41.4%; Pred. No. 7.1;  
Matches 12; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKQVEKALEQLEDDKVK 29  
Db 27 KAMEBKVKAMEBKVKAMEBKVKAMEBKVK 55

## RESULT 24

US-10-023-066A-58  
Sequence 58, Application US/10023066A  
Publication No. US20030056242A1  
GENERAL INFORMATION:

APPLICANT: E. I. DU PONT DE NEMOURS AND

COMPANY

TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR

INCREASING THE LYSINE AND

INCREASING THE LYSINE AND

THREONINE CONTENT OF THE SEEDS OF  
PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS

AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD VERSION 2.0C

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/023,066A

FILING DATE: 29-Apr-2002

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BARBARA C. SIEGEL

REGISTRATION NUMBER: 30,684

REFERENCE/DOCKET NUMBER: BB-1037-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-4931

TELEFAX: 302-773-0164

TELEX: 835420

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 28 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 58:

US-10-023-066A-58

Query Match 44.5%; Score 61; DB 14; Length 28;  
Best Local Similarity 42.3%; Pred. No. 3.2;  
Matches 11; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Qy 4 EDKVKASREAKQVEKALEQLEDDKVK 29  
Db 2 EKKKAMEBKVKAMEBKVKAMEBKVK 27

## RESULT 25

US-10-804-678-58  
Sequence 58, Application US/10804678  
Publication No. US20050005330A1  
GENERAL INFORMATION:

APPLICANT: EPELBAUM, SABINE URSULA

FALCO, SAVERIO CARL

MCDEVITT, RAYMOND ERVIN, III

TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR

INCREASING THE LYSINE CONTENT OF

THE SEEDS OF PLANTS

NUMBER OF SEQUENCES: 132

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: MICROSOFT OFFICE 97

SOFTWARE: MICROSOFT WINDOWS 95

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/804,678

FILING DATE: 19-Mar-2004



CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/049,304  
FILING DATE: 27-Mar-1998  
APPLICATION NUMBER: 08/824,627  
FILING DATE: MARCH 27, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: CHRISTENDURY, LYNN M.  
REGISTRATION NUMBER: 30,971  
REFERENCE/DOCKET NUMBER: BB-1037-F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-5481  
TELEFAX: 302-892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 58:  
US-10-804-678-58

Query Match 44.5%; Score 61; DB 17; Length 28;  
Best Local Similarity 42.3%; Pred. No. 3.2;  
Matches 11; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Oy 4 EDKVKASREAKQVKEKLEOLEDKVK 29  
Db 2 EKMKAMEKMKMEKMKLEKMKERK 27

RESULT 26  
US-10-023-066A-85  
Sequence 85, Application US/10023066A  
Publication No. US20030056242A1  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR INCREASING THE LYSINE AND THREONINE CONTENT OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,066A  
FILING DATE: 29-Apr-2002  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIEGEL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 85:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 85:  
US-10-023-066A-85

Query Match 44.5%; Score 61; DB 14; Length 37;  
Best Local Similarity 37.9%; Pred. No. 4.2;  
Matches 11; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Oy 1 KQAEKVKASREAKQVKEKLEOLEDKVK 29  
Db 6 KLEEKMKAMEKMKMEKMKLEKMKERK 34

RESULT 27  
US-10-023-066A-86  
Sequence 86, Application US/10023066A  
Publication No. US20030056242A1  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR INCREASING THE LYSINE AND THREONINE CONTENT OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,066A  
FILING DATE: 29-Apr-2002  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIEGEL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 86:  
US-10-023-066A-86

Query Match 44.5%; Score 61; DB 14; Length 37;  
Best Local Similarity 37.9%; Pred. No. 4.2;  
Matches 11; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Oy 1 KQAEKVKASREAKQVKEKLEOLEDKVK 29  
Db 6 KLEEKMKAMEKMKMEKMKLEKMKERK 34

RESULT 28  
US-10-804-678-85  
Sequence 85, Application US/10804678  
Publication No. US20050005330A1



REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 56 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 77:  
US-10-023-066A-77

Query Match 44.5%; Score 61; DB 14; Length 56;  
Best Local Similarity 41.4%; Pred. No. 6.6;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KOAEKVKASREAKQVEKALEQLEDKVK 29  
Db 6 KAMEKMKAMEKMKAMEKMKAMEKMK 34

RESULT 31  
US-10-804-678-77  
Sequence 77, Application US/10804678  
Publication No. US200500530A1  
GENERAL INFORMATION:  
APPLICANT: EPELBAUM, SABINE URSULA  
FALCO, SAVERIO CARL  
MCDEVITT, RAYMOND ERVIN, III  
TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
INCREASING THE LYSINE CONTENT OF  
THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT OFFICE 97  
SOFTWARE: MICROSOFT WINDOWS 95  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/804,678  
FILING DATE: 19-Mar-2004  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/049,304  
FILING DATE: 27-Mar-1998  
APPLICATION NUMBER: 08/824,627  
FILING DATE: MARCH 27, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: CHRISTENBURY, LYNN M.  
REGISTRATION NUMBER: 30,971  
REFERENCE/DOCKET NUMBER: BB-1037-F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-5481  
TELEFAX: 302-892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 56 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 77:  
US-10-804-678-77

Query Match 44.5%; Score 61; DB 17; Length 56;  
Best Local Similarity 41.4%; Pred. No. 6.6;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KOAEKVKASREAKQVEKALEQLEDKVK 29  
Db 6 KAMEKMKAMEKMKAMEKMKAMEKMK 34

RESULT 32  
US-10-437-963-146357  
Sequence 146357, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovallik, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 146357  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_4699C.1.pep  
US-10-437-963-146357

Query Match 44.5%; Score 61; DB 16; Length 145;  
Best Local Similarity 48.3%; Pred. No. 18;  
Matches 14; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 KOAEKVKASREAKQVEKALEQLEDKVK 29  
Db 80 KEAEKKKKQDEBEKKKEKEKEBEKKKK 108

RESULT 33  
US-10-437-963-146368  
Sequence 146368, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovallik, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 146368  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_4699C.1.pep  
US-10-437-963-146368

Query Match 44.5%; Score 61; DB 16; Length 145;

Best Local Similarity 48.3%; Pred. No. 18;  
Matches 14; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 KQADKYASREAKQVEKALEQLEDDYK 29  
Db 80 KEAEKKKKDEBEKKKKKEKEKEBEKKKK 108

## RESULT 34

US-10-023-066A-62  
; Sequence 62, Application US/10023066A  
; Publication No. US20030056242A1  
; GENERAL INFORMATION:  
; APPLICANT: E. I. DU PONT DE NEMOURS AND  
; COMPANY

TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
INCREASING THE LYSINE AND  
THREONINE CONTENT OF THE SEEDS OF  
PLANTS

NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS

STREET: 1007 MARKET STREET  
CITY: WILMINGTON

STATE: DELAWARE  
COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/023,066A  
FILING DATE: 29-Apr-2002

CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:

NAME: BARBARA C. SIEGEL  
REGISTRATION NUMBER: 30,684

REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164

TELEX: 835420  
INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids

TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-10-023-066A-62

Query Match 43.8%; Score 60; DB 14; Length 35;  
Best Local Similarity 38.5%; Pred. No. 5.2;  
Matches 10; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Qy 4 EDKYASREAKQVEKALEQLEDDYK 29  
Db 2 EKKKAMEBKMKKEKKKKMEKKKK 27

US-10-023-066A-62

RESULT 35

US-10-804-678-62

; Sequence 62, Application US/10804678  
; Publication No. US2005005330A1  
; GENERAL INFORMATION:  
; APPLICANT: EPELBAUM, SABINE URSULA  
; PALCO, SAVERIO CARL  
; MCDEVITT, RAYMOND BRYAN, III  
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
; INCREASING THE LYSINE CONTENT OF

THE SEEDS OF PLANTS

NUMBER OF SEQUENCES: 132

CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

STREET: 1007 MARKET STREET  
CITY: WILMINGTON

STATE: DELAWARE  
COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT OFFICE 97

SOFTWARE: MICROSOFT WINDOWS 95  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/804,678  
FILING DATE: 19-Mar-2004

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/049,304  
FILING DATE: 27-Mar-1998

APPLICATION NUMBER: 08/824,627  
FILING DATE: MARCH 27, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: CHRISTENBURY, LYNN M.

REGISTRATION NUMBER: 30,971  
REFERENCE/DOCKET NUMBER: BB-1037-F

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-5481

TELEFAX: 302-892-7949  
TELEX: 835420

INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:

LENGTH: 35 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 62:  
US-10-804-678-62

Query Match 43.8%; Score 60; DB 17; Length 35;  
Best Local Similarity 38.5%; Pred. No. 5.2;  
Matches 10; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Qy 4 EDKYASREAKQVEKALEQLEDDYK 29  
Db 2 EKKKAMEBKMKKEKKKKMEKKKK 27

US-10-023-066A-34

US-10-023-066A-34

; Sequence 34, Application US/10023066A  
; Publication No. US20030056242A1  
; GENERAL INFORMATION:  
; APPLICANT: E. I. DU PONT DE NEMOURS AND  
; COMPANY

TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
INCREASING THE LYSINE AND  
THREONINE CONTENT OF THE SEEDS OF  
PLANTS

NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS

STREET: 1007 MARKET STREET  
CITY: WILMINGTON

STATE: DELAWARE  
COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62

US-10-023-066A-62



US-10-023-066A-32  
; Sequence 32, Application US/10023066A  
; Publication No. US20030056242A1  
; GENERAL INFORMATION:  
; APPLICANT: E. I. DU PONT DE NEMOURS AND  
; COMPANY  
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
; INCREASING THE LYSINE AND  
; THREONINE CONTENT OF THE SEEDS OF  
; PLANTS  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS  
; AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: U.S.A.  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MICROSOFT WORD VERSION 2.0C  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,066A  
; FILING DATE: 29-Apr-2002  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BARBARA C. SIEGELL  
; REGISTRATION NUMBER: 30,684  
; REFERENCE/DOCKET NUMBER: BB-1037-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-992-4931  
; TELEFAX: 302-773-0164  
; TELEX: 835420  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 49 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-10-023-066A-32  
Query Match 43.8%; Score 60; DB 14; Length 49;  
Best Local Similarity 41.4%; Pred. No. 7.4;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
QY 1 KOAEDKVKASREAKQVEKALEQLEDKVK 29  
| : : : | : : : | : : : | : : : | : : : |  
| : : : | : : : | : : : | : : : | : : : |  
DB 6 KAMEEKLKAMEEKLKAMEEKLKAMEEKLK 34  
RESULT 40  
US-10-023-066A-54  
; Sequence 54, Application US/10023066A  
; Publication No. US20030056242A1  
; GENERAL INFORMATION:  
; APPLICANT: E. I. DU PONT DE NEMOURS AND  
; COMPANY  
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
; INCREASING THE LYSINE AND  
; THREONINE CONTENT OF THE SEEDS OF  
; PLANTS  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS  
; AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: U.S.A.

ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MICROSOFT WORD VERSION 2.0C  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,066A  
; FILING DATE: 29-Apr-2002  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BARBARA C. SIEGELL  
; REGISTRATION NUMBER: 30,684  
; REFERENCE/DOCKET NUMBER: BB-1037-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-992-4931  
; TELEFAX: 302-773-0164  
; TELEX: 835420  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 49 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 54:  
US-10-023-066A-54  
Query Match 43.8%; Score 60; DB 14; Length 49;  
Best Local Similarity 41.4%; Pred. No. 7.4;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
QY 1 KOAEDKVKASREAKQVEKALEQLEDKVK 29  
| : : : | : : : | : : : | : : : | : : : |  
| : : : | : : : | : : : | : : : | : : : |  
DB 6 KAMEEKLKAMEEKLKAMEEKLKAMEEKLK 34

Search completed: June 13, 2005, 21:03:34  
Job time : 141.628 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: June 13, 2005, 20:28:35 ; Search time 32.3721 Seconds  
(without alignment)  
86.194 Million cell updates/sec

Title: US-10-706-275-2

Perfect score: 137  
Sequence: 1 KOAEDKVKASREAKQVEKALEQEDKVK 29Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72.5	52.9	587	2 JCI1419	FC gamma (IgG) rec
2	72	52.6	388	2 A49545	plasmnogen-bindin
3	72	52.6	408	2 S30283	protein M precurs
4	72	52.6	436	2 S30284	M protein precurs
5	72	52.6	454	2 S43556	plasmnogen-bindin
6	72	52.6	472	2 S43554	plasmnogen-bindin
7	72	52.6	483	2 A26297	M6 protein - Strept
8	72	52.6	484	2 S35401	M1 protein precurs
9	72	52.6	484	2 S46469	M1 protein precurs
10	72	52.6	484	2 S34978	M1 protein precurs
11	72	52.6	501	2 A44643	M protein precurs
12	72	52.6	532	2 S54871	M protein - Strept
13	72	52.6	564	2 A60115	M protein precurs
14	71	51.8	492	2 A28616	M5 protein precurs
15	71	51.8	539	2 A28549	M24 protein precurs
16	70	51.1	528	2 S57835	IgG-binding protei
17	66	48.2	1365	2 T45031	hypothetical prote
18	66	48.2	1408	2 T45039	hypothetical prote
19	64	46.7	217	2 G75219	hypothetical prote
20	63.5	46.4	104	1 H64337	conserved hypotnet
21	62	45.3	493	2 T21260	hypothetical prote
22	61	44.5	445	2 T50972	probable zuotin [i
23	60	43.8	284	2 S23470	alanine-tropomyosin
24	59	43.1	876	2 AFO843	beta-tropomyosin -
25	59	43.1	284	1 TMRRB	tropomyosin beta c
26	59	43.1	284	1 TMRRB	tropomyosin 1, emb
27	59	43.1	284	1 A23562	tropomyosin 1, fib
28	59	43.1	284	2 S00922	tropomyosin beta,
29	59	43.1	284	2 A44131	tropomyosin beta 2

30	59	43.1	284	2 S03838	tropomyosin beta,
31	59	43.1	284	2 S23256	tropomyosin beta -
32	59	43.1	284	2 B25073	tropomyosin beta,
33	59	43.1	436	2 AH1387	cell wall binding
34	58	42.3	217	2 B71203	hypothetical prote
35	58	42.3	284	1 TWCHS1	tropomyosin 1, smo
36	58	42.3	284	2 A30125	tropomyosin 1, ske
37	58	42.3	751	2 T34490	hypothetical prote
38	58	42.3	2116	2 A26655	myosin heavy chain
39	57.5	42.0	558	2 D70449	conserved hypotnet
40	57	41.6	169	2 E97357	hypothetical prote
41	57	41.6	174	2 A11274	general stress pro
42	57	41.6	248	2 A25530	tropomyosin, fibro
43	57	41.6	248	2 S11390	tropomyosin 5 - mo
44	57	41.6	248	2 S34124	tropomyosin isoform
45	57	41.6	248	2 I53784	tropomyosin - rat

## ALIGNMENTS

```
RESULT 1
JCI1419
FC gamma (IgG) receptor II precursor - Streptococcus sp.
N:Alternate names: fcrv protein
C:Species: Streptococcus sp.
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: JCI1419, S17354
R:Smirnov, O.Y.; Denesnyk, A.I.; Zakharov, M.V.; Abramov, V.M.; Zav'yalov, V.P.
Gene 120, 27-32, 1992
A:Title: Protein V, a novel type-II IgG receptor from Streptococcus sp.: Sequence, homol
A:Reference number: JCI1419; PMID:93013016; PMID:1398120
A:Accession: JCI1419
A:Molecule type: DNA
A:Residues: 1-587 <SMT>
A:Cross-references: UNIPROT:Q55312; EXBL:X62467; NID:947562; PIDN:CAA44324.1; PID:947563
A:Experimental source: strain 22/58'Valente'
C:Genetics:
A:Gene: fcrv
C:Superfamily: M5 protein
C:Keywords: duplication; immunoglobulin receptor
F:1-41/Domain: signal sequence #status predicted <SIG>
F:42-587/Product: IgG FC receptor II #status predicted <MAT>
F:234-268/Region: 35-residue repeat A
F:269-303/Region: 35-residue repeat A
F:304-338/Region: 35-residue repeat A
F:339-373/Region: 35-residue repeat A
F:374-408/Region: 35-residue repeat B
F:416-450/Region: 35-residue repeat B

Query Match 52.9%; Score 72.5; DB 2; Length 587;
Best Local Similarity 46.3%; Pred. No. 2.5;
Matches 19; Conservative 3; Mismatches 6; Indels 13; Gaps 1;

QY 1 KOAEDK-----VASREAKQVEKALEQEDKVK 28
Db 417 KVKEKQISDASRQGLRDLDAASREAKQVEKALEFANSKL 457

RESULT 2
A49545
Plasmnogen-binding protein PAM precursor - Streptococcus pyogenes (fragment)
N:Alternate names: plasmnogen-binding M-like protein (Pd 53)
C:Species: Streptococcus pyogenes
C>Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: A49545; S61084; S60829; S70459; S32619
R:Berge, A.; Sjoebyring, U.
J. Biol. Chem. 268, 25417-25424, 1993
A:Title: PAM, a novel plasmnogen-binding protein from Streptococcus pyogenes.
A:Reference number: A49545; PMID:94064605; PMID:8244975
A:Accession: A49545
A:Molecule type: DNA
A:Residues: 1-388 <BER>
```

A/Cross-references: UNIPROT:P49054; EMBL:Z22219; NID:G288978; PIDN:CAA60222.1; PID:G9408  
R;Matmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
submitted to the EMBL Data Library, July 1994  
A/Description: Noncongruent relationships between variation in emm1 gene sequences and t  
A/Reference number: S61084  
A/Accession: S61084  
A/Molecule type: DNA  
A/Residues: 13-96 <MHA>  
A/Cross-references: EMBL:U11975; NID:G533627; PIDN:AAA9591.1; PID:G1235829  
A/Experimental source: serotype M53  
R;Matmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A/Title: Non-congruent relationships between variation in emm gene sequences and the po  
A/Reference number: S60784; MUID:95198537; PMID:7891551  
A/Accession: S60829  
A/Status: nucleic acid sequence not shown  
A/Molecule type: DNA  
A/Residues: 17-77 <MHA>  
A/Cross-references: EMBL:U11975  
A/Experimental source: serotype M53  
R;Carlsson Wistedt, A.; Ringdahl, U.; Mueller-Esterl, W.; Sjoebing, U.  
Mol. Microbiol. 18, 569-578, 1995  
A/Title: Identification of a plasminogen-binding motif in PAM, a bacterial surface prote  
A/Reference number: S70457; MUID:96342385; PMID:8748039  
A/Accession: S70459  
A/Molecule type: DNA  
A/Residues: 30-162 <CAR>  
C/Superfamily: M5 protein  
C/Keywords: cell wall  
F;1-29/Domain: signal sequence (fragment) #status predicted <SIG>  
F;30-388/Product: plasminogen-binding protein PAM #status predicted <MAT>

Query Match 52.6%; Score 72; DB 2; Length 388;  
Best Local Similarity 45.2%; Pred. No. 1.9;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLDEKV 28  
Db 244 DKVBEKQISDASRQGLRDLDSREAKQVEKALEANSKL 285

RESULT 3  
S30283  
protein M precursor - Streptococcus pyogenes (serotype M41)  
C/Species: Streptococcus pyogenes  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C/Accession: S30283; S29680  
R;Podbielski, A.  
Mol. Gen. Genet. 237, 287-300, 1993  
A/Title: Three different types of organization of the vir regulon in group A streptococ  
A/Reference number: S30283; MUID:93204905; PMID:8455563  
A/Accession: S30283  
A/Status: nucleic acid sequence not shown  
A/Molecule type: DNA  
A/Residues: 1-408 <POD1>  
A/Cross-references: UNIPROT:Q54837; EMBL:X58178  
R;Podbielski, A.; Melzer, B.  
submitted to the EMBL Data Library, February 1991  
A/Description: Pcr mediated cloning and sequencing of group A streptococcal emm41/52 (ev  
A/Reference number: S29680  
A/Accession: S29680  
A/Molecule type: DNA  
A/Residues: 1-230, 'N', 233-371, 'R', 373-408 <POD2>  
A/Cross-references: EMBL:X58178; NID:G47362; PIDN:CAA41167.1; PID:G47363  
C/Genetics:  
A/Gene: emm  
C/Superfamily: M5 protein  
C/Keywords: transmembrane protein  
F;1-41/Domain: signal sequence #status predicted <SIG>  
F;42-408/Product: M protein #status predicted <MAT>  
F;383-401/Domain: transmembrane #status predicted <TMM>

Query Match 52.6%; Score 72; DB 2; Length 408;

Best Local Similarity 50.0%; Pred. No. 2;  
Matches 18; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

Qy 1 KQAE-----KVRASREAKQVEKALEQLDEKV 28  
Db 243 KQVSDASRQGLRDLDSREAKQVEKALEANSKL 278

RESULT 4  
S30284  
M protein precursor - Streptococcus pyogenes (serotype M52)  
C/Species: Streptococcus pyogenes  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C/Accession: S30284; S29681  
R;Podbielski, A.  
Mol. Gen. Genet. 237, 287-300, 1993  
A/Title: Three different types of organization of the vir regulon in group A streptococci  
A/Reference number: S30283; MUID:93204905; PMID:8455563  
A/Accession: S30284  
A/Status: nucleic acid sequence not shown  
A/Molecule type: DNA  
A/Residues: 1-436 <POD1>  
A/Cross-references: UNIPROT:Q54839; EMBL:X58179  
R;Podbielski, A.; Melzer, B.  
submitted to the EMBL Data Library, February 1991  
A/Description: Pcr mediated cloning and sequencing of group A streptococcal emm41/52 (ev  
A/Reference number: S29680  
A/Accession: S29681  
A/Molecule type: DNA  
A/Residues: 1-216, 'N', 218-436 <POD2>  
A/Cross-references: EMBL:X58179; NID:G47364; PIDN:CAA41168.1; PID:G47365  
C/Genetics:  
A/Gene: emm  
C/Superfamily: M5 protein  
C/Keywords: transmembrane protein  
F;1-41/Domain: signal sequence #status predicted <SIG>  
F;42-436/Product: M protein #status predicted <MAT>  
F;411-429/Domain: transmembrane #status predicted <TMM>

Query Match 52.6%; Score 72; DB 2; Length 436;  
Best Local Similarity 45.2%; Pred. No. 2.1;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLDEKV 28  
Db 265 DKVBEKQISDASRQGLRDLDSREAKQVEKALEANSKL 306

RESULT 5  
S43596  
plasminogen-binding protein MLC36 - Streptococcus sp. (fragment)  
C/Species: Streptococcus sp.  
C/Date: 14-Sep-1994 #sequence\_revision 03-Nov-1995 #text\_change 26-Aug-1999  
C/Accession: S43598; S43596  
R;ben Nasr, A.; Wistedt, A.; Ringdahl, U.; Sjoebing, U.  
Eur. J. Biochem. 222, 267-276, 1994  
A/Title: Streptokinase activates plasminogen bound to human group C and G streptococci  
A/Reference number: S45598; MUID:94291620; PMID:8020466  
A/Accession: S45598  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-454 <BE2>  
A/Cross-references: EMBL:Z32677; NID:G474767; PIDN:CAA83588.1; PID:G474768  
C/Superfamily: M5 protein

Query Match 52.6%; Score 72; DB 2; Length 454;  
Best Local Similarity 45.2%; Pred. No. 2.2;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLDEKV 28  
Db 310 DKVBEKQISDASRQGLRDLDSREAKQVEKALEANSKL 351



RESULT 6  
S3354  
A:plasmogen-binding protein M1G72 - Streptococcus sp. (fragment)  
C:Species: Streptococcus sp.  
A:Variety: group G  
C:Date: 07-Sep-1994 #sequence\_revision 03-Nov-1995 #text\_change 20-Jun-2000  
C:Accession: S45599; S43554  
R:Ben Naser, A.; Wistedt, A.; Ringdahl, U.; Sjoeborg, U.  
Eur. J. Biochem. 222, 267-276, 1994  
A:Title: Streptokinase activates plasminogen bound to human group C and G streptococci  
A:Reference number: S45599; MUID:94291620; PMID:8020466  
A:Accession: S45599  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-472 <BE2>  
A:Cross-references: EMBL:Z32678; NID:9474769; PIDN:CAA83589.1; PID:g133383  
C:Superfamily: M5 protein

Query Match 52.6%; Score 72; DB 2; Length 472;  
Best Local Similarity 45.2%; Pred. No. 2.3;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 328 DKVKEKQISDTSRKGLRDLDSREAKQVEKALEANSKL 369

RESULT 7  
A26297  
M6 protein - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 09-Jul-2004  
C:Accession: A26297  
R:Hollingshead, S. K.; Fischetti, V. A.; Scott, J. R.  
J. Biol. Chem. 261, 1677-1686, 1986  
A:Title: Complete nucleotide sequence of type 6 M protein of the group A streptococcus.  
A:Reference number: A26297; MUID:86111835; PMID:3511046  
A:Accession: A26297  
A:Molecule type: DNA  
A:Residues: 1-483 <HOL>  
A:Cross-references: UNIPROT:P08089; GB:M11338; GB:M11415; NID:g153699; PIDN:AAA26920.1;  
C:Genetics: emme  
A:Gene: emme  
C:Superfamily: M5 protein  
C:Keywords: coiled coil; transmembrane protein

Query Match 52.6%; Score 72; DB 2; Length 483;  
Best Local Similarity 45.2%; Pred. No. 2.4;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 312 DKVKEKQISDASRQGLRDLDSREAKQVEKALEANSKL 353

RESULT 8  
S35401  
M1 protein precursor - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
A:Variety: serotype M1  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S35401; S61074; S60784  
R:Podileski, A.  
submitted to the EMBL Data Library, September 1991  
A:Reference number: S35401  
A:Accession: S35401  
A:Molecule type: DNA  
A:Residues: 1-484 <PDB>  
A:Cross-references: UNIPROT:Q10372; EMBL:X62131; NID:g311757; PIDN:CAA44062.1; PID:g3117  
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
submitted to the EMBL Data Library, July 1994  
A:Title: Noncongruent relationships between variation in emm1 gene sequences and v

A:Reference number: S61072  
A:Accession: S61074  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 16-94 <MHA>  
A:Cross-references: EMBL:U11940; NID:9533557; PIDN:AAA99556.1; PID:9533558  
R:Whitmore, A. M.; Kapur, V.; Sullivan, D. J.; Musser, J. M.; Kehoe, M. A.  
Mol. Microbiol. 14, 619-631, 1994  
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop.  
A:Reference number: S60784; MUID:95196537; PMID:7891551  
A:Accession: S60784  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 29-89 <MH2>  
A:Cross-references: EMBL:U11940  
C:Genetics: emm1  
A:Gene: emm1  
C:Superfamily: M5 protein

Query Match 52.6%; Score 72; DB 2; Length 484;  
Best Local Similarity 45.2%; Pred. No. 2.4;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 313 DKVKEKQISDASRQGLRDLDSREAKQVEKALEANSKL 354

RESULT 9  
S46489  
M1 protein precursor - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S46489; S46490  
R:Axelson, P.; Schmidt, K. H.; Cooney, J.; Björck, L.  
Biochem. J. 300, 877-886, 1994  
A:Title: M1 protein and protein H: IgGFC- and albumin-binding streptococcal surface prot.  
A:Reference number: S46489; MUID:94280417; PMID:8010973  
A:Accession: S46489  
A:Molecule type: DNA  
A:Residues: 1-484 <AKE>  
A:Cross-references: UNIPROT:Q05464; UNIPROT:Q10372; UNIPROT:Q09XV0  
A:Experimental source: strain 40/58, serotype M1  
A:Accession: S46490  
A:Molecule type: protein  
A:Residues: 42-51 <AKW>  
A:Experimental source: strain 40/58, serotype M1  
C:Genetics: emm1  
A:Gene: emm1  
C:Superfamily: M5 protein  
C:Keywords: transmembrane protein  
P:1-41/Domain: signal sequence #status predicted <SIG>  
P:42-484/Product: M1 protein #status experimental <MAT>  
P:459-477/Domain: transmembrane #status predicted <TM>

Query Match 52.6%; Score 72; DB 2; Length 484;  
Best Local Similarity 45.2%; Pred. No. 2.4;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 313 DKVKEKQISDASRQGLRDLDSREAKQVEKALEANSKL 354

RESULT 10  
S34978  
M1.1 protein precursor - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S34978; S31966  
R:Harbaugh, M. P.; Podileski, A.; Huegl, S.; Cleary, P. P.  
Mol. Microbiol. 8, 981-991, 1993  
A:Title: Nucleotide substitutions and small-scale insertion produce size and antigenic v.

A:Reference number: S34978; MUID:93360826; PMID:8355619  
A:Accession: S34978  
A:Molecule type: DNA  
A:Residues: 1-484 <HAR>  
A:Cross-references: UNIPROT:005464; EMBL:Z21845; NID:g49401; PIDN:CAA79893.1; PID:g49402  
C:Genetics:  
A:Gene: emm1.1  
C:Superfamily: M5 protein  
C:Keywords: membrane protein  
P:1-42/Domain: signal sequence #status predicted <SIG>  
F:43-484/Product: M1.1 protein #status predicted <MAT>

Query Match 52.6%; Score 72; DB 2; Length 484;  
Best Local Similarity 45.2%; Pred. No. 2.4;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

OY 5 DKVK-----ASREAKKQVEKALEQLEDKV 28  
Db 313 DKVKEKQISDASRQGLRRDLDSREAKKQVEKALEANSKL 354

RESULT 11  
A44643  
M protein precursor - Streptococcus pyogenes (serotype M57) (fragment)  
C:Species: Streptococcus pyogenes  
A:Variety: serotype M57  
C:Date: 26-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
R:Manjula, B.N.; Khandke, K.M.; Fairwell, T.; Reif, W.A.; Sriprakash, K.S.  
J: Protein Chem. 10, 369-384, 1991  
A:Title: Heptad motifs within the distal subdomain of the coiled-coil rod region of M protein: nucleotide sequence of the M57 gene and relation of the dedu.  
A:Reference number: A44643; MUID:92143933; PMID:1781883  
A:Accession: A44643  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-501 <MAN>  
A:Experimental source: type M57, strain A595  
A>Note: sequence inconsistent with nucleotide translation  
A>Note: sequence extracted from NCBI backbone (NCBIN:83737, NCBI:P:83738)  
A>Note: parts of this sequence were confirmed by peptide sequencing  
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop  
A:Reference number: S60784; MUID:95198537; PMID:7891551  
A:Accession: S60833  
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 14-95 <WHA>  
A:Cross-references: EMBL:U11971  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Superfamily: M5 protein  
C:Keywords: coiled coil; dimer

Query Match 52.6%; Score 72; DB 2; Length 501;  
Best Local Similarity 45.2%; Pred. No. 2.4;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

OY 5 DKVK-----ASREAKKQVEKALEQLEDKV 28  
Db 342 DKVKEKQISDASRQGLRRDLDSREAKKQVEKALEANSKL 383

RESULT 12  
S54871  
M protein - Streptococcus sp.  
C:Species: Streptococcus sp.  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
A:Accession: S54871  
R:Podbielski, A.; Melzer, B.  
Submitted to the EMBL Data Library, June 1991  
A:Reference number: S54871  
A:Accession: S54871

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-532 <POD>  
A:Cross-references: UNIPROT:Q55098; EMBL:X60097; NID:g840905; PIDN:CAA42693.1; PID:g84091  
C:Superfamily: M5 protein

Query Match 52.6%; Score 72; DB 2; Length 532;  
Best Local Similarity 45.2%; Pred. No. 2.6;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

OY 5 DKVK-----ASREAKKQVEKALEQLEDKV 28  
Db 361 DKVKEKQISDASRQGLRRDLDSREAKKQVEKALEANSKL 402

RESULT 13  
A60115  
M protein precursor - Streptococcus pyogenes (serotype M12) (fragment)  
C:Species: Streptococcus pyogenes  
A:Variety: serotype M12  
C:Date: 08-Dec-1992 #sequence\_revision 08-Dec-1992 #text\_change 09-Jul-2004  
A:Accession: A40174; A60115; S39887; S61072; S60793  
R:Robbins, J.C.; Spanier, J.G.; Jones, S.J.; Simpson, W.J.; Cleary, P.P.  
J: Bacteriol. 169, 5633-5640, 1987  
A:Title: Streptococcus pyogenes type 12 M protein gene regulation by upstream sequences.  
A:Reference number: A40174; MUID:88058777; PMID:2445730  
A:Accession: A40174  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-564 <ROB>  
A:Cross-references: UNIPROT:P19401; GB:M18269; NID:g153543; PIDN:AAA88573.1; PID:g153544  
R:Kraus, W.; Seyer, J.M.; Beachey, E.H.  
Infect. Immun. 57, 2457-2461, 1989  
A:Title: Vimentin-cross-reactive epitope of type 12 streptococcal M protein.  
A:Reference number: A60115; MUID:89307564; PMID:2473037  
A:Accession: A60115  
A:Molecule type: protein  
A:Residues: 42-54 <KRA>  
R:Chen, C.; Bormann, N.; Cleary, P.P.  
Mol. Gen. Genet. 241, 685-693, 1993  
A:Title: V1RR and M1Y are homologous trans-acting regulators of M protein and C5a peptid.  
A:Reference number: S39886; MUID:94088463; PMID:7505389  
A:Accession: S39887  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-15 <CHS>  
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.  
Submitted to the EMBL Data Library, July 1994  
A:Description: Noncongruent relationships between variation in emm1 gene sequences and t  
A:Reference number: S61072  
A:Accession: S61072  
A:Molecule type: DNA  
A:Residues: 13-111 <WHA>  
A:Cross-references: EMBL:U11937; NID:g533551; PIDN:AAA99553.1; PID:g1235807  
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
Mol. Microbiol. 14, 619-631, 1994  
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop  
A:Reference number: S60784; MUID:95198537; PMID:7891551  
A:Accession: S60793  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 29-89 <WHW>  
A:Cross-references: EMBL:U11937  
C:Genetics:  
A:Gene: emm12  
C:Superfamily: M5 protein  
C:Keywords: transmembrane protein  
P:1-41/Domain: signal sequence #status predicted <SIG>  
F:42-564/Product: M protein (fragment) #status predicted <MAT>

Query Match 52.6%; Score 72; DB 2; Length 564;  
Best Local Similarity 45.2%; Pred. No. 2.7;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKKQVEKALEQLEDRV 28  
 |||||  
 Db 405 DKVKEBKQISDASRGGLRDLDASREAKKQVEKALEANSKL 446

## RESULT 14

A28616  
 M5 protein precursor - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes  
 A:Variety: serotype M5  
 C:Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 09-Jul-2004  
 A:Accession: A28616; S60787  
 R:Miller, L.; Gray, L.; Beachey, E.; Kehoe, M.  
 J. Biol. Chem. 263, 5668-5673, 1988  
 A:Title: Antigenic variation among group A streptococcal M proteins. Nucleotide sequence  
 A:Reference number: A28616; MUID:86186881; PMID:3281944  
 A:Accession: A28616  
 A:Molecule type: DNA  
 A:Residues: 1-492 <MTL>  
 A:Cross-references: UNIPROT:P02977; GB:M20374; NID:G153812; PIDN:AAA26976.1; PID:G153813  
 R:Wharmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
 Mol. Microbiol. 14, 619-631, 1994  
 A:Title: Non-congruent relationships between variation in emm gene sequences and the pof  
 A:Reference number: S60784; MUID:95198537; PMID:7891551  
 A:Accession: S60787  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: DNA  
 A:Residues: 30-89 <WMA>  
 C:Genetics:  
 A:Gene: emp5  
 C:Superfamily: M5 protein  
 C:Keywords: coiled coil; transmembrane protein  
 F:1-42/Domain: signal sequence #status predicted <SIG>  
 F:43-492/Product: M5 protein #status predicted <MAT>

Query Match 51.8%; Score 71; DB 2; Length 492;  
 Best Local Similarity 75.0%; Pred. No. 3;  
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ASREAKKQVEKALEQLEDRV 28  
 |||||  
 Db 343 ASREAKKQVEKALEANSKL 362

## RESULT 15

A28549  
 M24 protein precursor - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes  
 A:Variety: serotype M24  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
 A:Accession: A28549; S60802  
 R:Moew, A.R.; Beachey, E.H.; Burdett, V.  
 J. Bacteriol. 170, 676-684, 1988  
 A:Title: Molecular evolution of streptococcal M protein: cloning and nucleotide sequence  
 A:Reference number: A28549; MUID:88115166; PMID:3276665  
 A:Accession: A28549  
 A:Molecule type: DNA  
 A:Residues: 1-539 <MOU>  
 A:Cross-references: UNIPROT:P12379; GB:M19031; NID:G153616; PIDN:AAA26874.1; PID:G153617  
 R:Wharmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.  
 Mol. Microbiol. 14, 619-631, 1994  
 A:Title: Non-congruent relationships between variation in emm gene sequences and the pof  
 A:Reference number: S60784; MUID:95198537; PMID:7891551  
 A:Accession: S60802  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: DNA  
 A:Residues: 30-89 <WMA>  
 C:Superfamily: M5 protein  
 C:Keywords: coiled coil; transmembrane protein  
 Query Match 51.8%; Score 71; DB 2; Length 539;  
 Best Local Similarity 75.0%; Pred. No. 3.3;

Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 9 ASREAKKQVEKALEQLEDRV 28  
 |||||  
 Db 390 ASREAKKQVEKALEANSKL 409

## RESULT 16

S57835  
 Igg-binding protein emml precursor - Streptococcus pyogenes (strain 64/14)  
 N:Alternate names: Igc-binding protein type IIA; type IIA immunoglobulin G-binding prote  
 C:Species: Streptococcus pyogenes  
 A:Variety: strain 64/14  
 C:Date: 28-Nov-1995 #sequence\_revision 09-Mar-1996 #text\_change 09-Jul-2004  
 A:Accession: S57835; S58931  
 R:Boyle, M.D.P.; Hawlitzky, J.; Raeder, R.; Podbielski, A.  
 Infect. Immun. 62, 1336-1347, 1994  
 A:Title: Analysis of genes encoding two unique type IIA immunoglobulin G-binding protein  
 A:Reference number: S57834; MUID:94178942; PMID:8132341  
 A:Accession: S57835  
 A:Molecule type: DNA  
 A:Residues: 1-528 <BOY>  
 A:Cross-references: UNIPROT:Q54843; EMBL:X72932  
 A:Experimental source: strain 64/14  
 A>Note: the authors translated the codons CTGAA for residue 52 and 53 as Arg  
 R:Podbielski, A.  
 submitted to the EMBL Data Library, March 1993  
 A:Reference number: S58931  
 A:Accession: S58931  
 A:Molecule type: DNA  
 A:Residues: 1-46, 'E', 48-52, 'E', 54-528 <POD>  
 A:Cross-references: EMBL:X72932; NID:G507128; PIDN:CAA51437.1; PID:G507130  
 A:Experimental source: strain 64/14  
 C:Genetics:  
 A:Gene: emml  
 C:Superfamily: M5 protein  
 F:1-41/Domain: signal sequence #status predicted <SIG>  
 F:42-528/Product: type IIA immunoglobulin G-binding protein emml #status predicted <MAT>

Query Match 51.1%; Score 70; DB 2; Length 528;  
 Best Local Similarity 45.2%; Pred. No. 4.1;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKKQVEKALEQLEDRV 28  
 |||||  
 Db 357 DKVKEBKQISDASRGGLRDLDASREAKKQVEKALEANSKL 398

## RESULT 17

T45031  
 hypocholesterol protein Y39B6 e [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 A:Accession: T45031  
 R:Williamson, R.; Aincough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johns Nature 368, 32-38, 1994  
 A:Authors: Showkhen, R.; Sims, M.; Snaidon, N.; Smith, A.; Smith, M.; Sonhammer, E.; S lock, L.; Wilkinson-Sprat, J.; Wohlman, P.  
 A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.  
 A:Reference number: S43531; MUID:94150718; PMID:7906398  
 A:Accession: T45031  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1365 <WTL>  
 A:Cross-references: EMBL:AL32896; NID:96434440; PIDN:CAB60910.1; PID:G6434445  
 A:Experimental source: clone Y39B6  
 C:Genetics:  
 A:Map position: 3  
 A:Intons: 10/1; 37/1; 171/2; 844/2; 1074/2; 1115/3; 1194/2; 1233/3  
 A>Note: Y39B6B.e

Query Match 48.2% Score 66; DB 2; Length 1365;  
Best Local Similarity 48.3% Pred. No. 25;  
Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 1 KOAEDKVASREAKKQVEKALEQLEDKVK 29  
998 KEEBERVAKEREKVKKEERLKAEEKVK 926

RESULT 18  
T45039  
hypochemical protein Y3986B.m [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T45039  
R:Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; Fothergill, J.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnson, B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D. Nature 368, 32-38, 1994  
A:Authors: Shownkeen, R.; Sims, M.; Smaiden, N.; Smith, A.; Smith, M.; Sonnenhammer, E.; Slock, L.; Wilkinson-Spott, U.; Woldman, P.  
A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.  
A:Reference number: S43531; MUID:94150718; PMID:7906398  
A:Accession: T45039  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1408 <WTL>  
A:Cross-references: EMBL:AL132896; NID:G6434440; PIDN:CAB60918.1; PID:G6434453  
A:Experimental source: clone Y3986B  
C:Genetics:  
A:Map position: 3  
A:Introns: 10/; 37/1; 856/2; 1107/2; 1148/3; 1233/2; 1272/3  
A>Note: Y3986B.m

Query Match 48.2% Score 66; DB 2; Length 1408;  
Best Local Similarity 48.3% Pred. No. 26;  
Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 1 KOAEDKVASREAKKQVEKALEQLEDKVK 29  
910 KEEBERVAKEREKVKKEERLKAEEKVK 938

RESULT 19  
T45039  
hypochemical protein PAB2166 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: G75219  
R:anonymous, GenomeScope  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure submitted to the EMBL Data Library, July 1999  
A:Reference number: A75001  
A:Accession: G75219  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-217 <KAN>  
A:Cross-references: UNIPROT:Q9V1Z2; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CAB492020  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB2166

Query Match 46.7% Score 64; DB 2; Length 217;  
Best Local Similarity 37.9% Pred. No. 7;  
Matches 11; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

Db 1 KOAEDKVASREAKKQVEKALEQLEDKVK 29  
112 EKKEEIKRLREKEKIERELEELKRVK 140

RESULT 20  
H64327

```

conserved hypothetical protein MJ0223 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 10-Sep-1999 #sequence_revistion 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: H64327
R:Butt, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
J.; Reich, C.I.; Overbeek, R.; Kirsness, B.F.; Weissbrock, K.G.; Merrick, J.M.; Glodok, A.;
iron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MWID:196317999; PMID:868087
A:Accession: H64327
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-104 <BU>
A:Cross-references: UNIPROT:Q57676; GB:U67478; GB:L77117; NID:G1590958; PIDN:AAB98215.1;
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0223

Query Match 46.4%; Score 63.5; DB 1; Length 104;
Best Local Similarity 46.7%; Pred. No. 4.4;
Matches 14; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Oy 1 KOAEKV-KASREAKKQVEKALEQLEDPKVK 29
Db 54 KLVSEMTKKAEBEAKKEAKKILEETKEIK 83

RESULT 21
T22180
hypothetical protein F44Fl.6a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revistion 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22180
R:Dobson, R.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19527
A:Accession: T22180
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-493 <WIL>
A:Cross-references: UNIPROT:Q9U3F0; EMBL:Z81083; PIDN:CAB54248.1; GSPDB:GN000116; CESP:F4
A:Experimental source: clone F44Fl
C:Genetics:
A:Gene: CESP:F44Fl.6a
A:Map position: 1

Query Match 45.3%; Score 62; DB 2; Length 493;
Best Local Similarity 55.5%; Pred. No. 26;
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Oy 1 KOAEKV-KASREAKKQVEKALEQLEDPK 27
Db 466 KLEBAKVQASREAKKQYENIMTLIGEK 492

RESULT 22
T50972
probable znucrin [imported] - Neurospora crassa
N:Alternate names: protein B24P7.270
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revistion 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50972
R:Schlute, U.; Aism, V.; Hohenstein, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T50972
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-445 <SCH>
A:Cross-references: UNIPROT:Q9P3Q8; EMBL:AL389890; GSPDB:GN00116; NCSP:B24P7.270
A:Experimental source: BAC clone B24P7; strain OR74A

```

C:Genetics:  
A:Gene: NCSP:B24P7.270  
A:Map position: 6  
A:Introns: 98/3

Query Match 44.5%; Score 61; DB 2; Length 445;  
Best Local Similarity 44.8%; Pred. No. 30;  
Matches 13; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKKQVEKALEQLEDDKYK 29  
Db 332 KEASEKAKADREASKKAKKAAKNAVKKXK 360

## RESULT 23

beta-tropomyosin - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S23470  
R:Hardy, S.; Thiebaud, P.  
A:Title: Isolation and characterization of cDNA clones encoding the skeletal and smooth  
A:Reference number: S23470; MUID:92305070; PMID:1610908  
A:Accession: S23470  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-284 <HAR>  
A:Cross-references: UNIPROT:Q91706; EMBL:N87307; NID:g214001; PIDD:AA10100.1; PID:g2140  
A:Note: the authors translated the codon GAG for residue 14 as Asp  
C:Superfamily: tropomyosin

Query Match 43.8%; Score 60; DB 2; Length 284;  
Best Local Similarity 39.3%; Pred. No. 25;  
Matches 11; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKKQVEKALEQLEDDKY 28  
Db 233 KEASRVAFPAKSVVQLEKTIIDLEDEV 260

## RESULT 24

alanine-tRNA ligase (EC 6.1.1.7) - Salmonella enterica subsp. enterica serovar Typhi (str  
A:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 17-Mar-2003  
C:Accession: AF0843  
R:Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,  
th, T.; Connerston, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0843  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-876 <PAR>  
A:Cross-references: GB:AL513382; PIDD:CAD05933.1; PID:gl6503904; GSPDB:GN00176  
C:Genetic: STY2948  
A:Gene: STY2948  
C:Superfamily: alanyl-tRNA ligase  
C:Keywords: ligase

Query Match 43.8%; Score 60; DB 2; Length 876;  
Best Local Similarity 52.2%; Pred. No. 71;  
Matches 12; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 5 DKVKASREAKKQVEKALEQLEDDK 27  
Db 730 DKRAVAVERTROLEKELOOLKQ 752

## RESULT 25

tropomyosin beta chain, skeletal muscle [validated] - rabbit (tentative sequence)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Jul-1980 #sequence\_revision 31-Jul-1980 #text\_change 09-Jul-2004  
C:Accession: A02980; I46509  
R:MaK, A.S.; Smillie, L.B.; Steward, G.R.  
J. Biol. Chem. 255, 3647-3655, 1980  
A:Title: A comparison of the amino acid sequences of rabbit skeletal muscle alpha- and b  
A:Reference number: A02980; MUID:80159993; PMID:7364764  
A:Accession: A02980  
A:Molecule type: protein  
A:Residues: 1-284 <MAK>  
A:Cross-references: UNIPROT:P58776  
A:Note: peptides that appeared to be identical with corresponding peptides of tropomyosi  
A:Note: minor heterogeneity was detected at 11 positions  
R:Putney, S.D.; Herlihy, W.C.; Schimmel, P.  
Nature 302, 718-721, 1983  
A:Title: A new tropomyosin T and cDNA clones for 13 different muscle proteins, found by sho  
A:Reference number: I46471; MUID:83167564; PMID:6687628  
A:Accession: I46509  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 131-172 <PUT>  
A:Cross-references: EMBL:V00893; NID:g1728; PIDD:CAA24258.1; PID:g929763  
C:Superfamily: tropomyosin

C:Keywords: acetylated amino end; alternative splicing; coiled coil  
P1-284/Product: tropomyosin beta chain #status experimental <MAT>  
F11/Modified site: acetylated amino end (Met) #status experimental

Query Match 43.1%; Score 59; DB 1; Length 284;  
Best Local Similarity 41.4%; Pred. No. 32;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KOAEDKVKASREAKKQVEKALEQLEDDKYK 29  
Db 30 KOAEDRCQLEBEOQALQKLIKGTEDVE 58

## RESULT 26

tropomyosin 1, embryonic fibroblast - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: A02981; A25073  
R:Yamawaki-Kataoka, Y.; Helfman, D.M.  
J. Biol. Chem. 260, 14440-14445, 1985  
A:Title: Rat embryonic fibroblast tropomyosin 1. cDNA and complete primary amino acid se  
A:Reference number: A02981; MUID:86033945; PMID:3840484  
A:Accession: A02981  
A:Molecule type: mRNA  
A:Residues: 1-284 <YAM>  
A:Cross-references: UNIPROT:P58775  
R:Helfman, D.M.; Cheley, S.; Kuismanen, E.; Finn, L.A.; Yamawaki-Kataoka, Y.  
Mol. Cell. Biol. 6, 3582-3595, 1986  
A:Title: Nonmuscle and muscle tropomyosin isoforms are expressed from a single gene by a  
A:Reference number: A93081; MUID:87089698; PMID:2433392  
A:Accession: A25073  
A:Molecule type: mRNA  
A:Residues: 1-284 <HEL>  
A:Cross-references: GB:L00382; GB:M14127; NID:g207495; PIDD:AAA42289.1; PID:g207498  
C:Comment: Rat embryonic fibroblasts contain three major (1, 2, and 4) and two minor (3  
dicates that they belong to the same isoform.  
C:Comment: The presence of TM in cells devoid of tropomyosin (nonmuscle and smooth muscle c  
s of these cell types, unlike skeletal and cardiac muscle TM, do not appear to be subjec  
C:Superfamily: tropomyosin  
C:Keywords: alternative splicing; coiled coil

Query Match 43.1%; Score 59; DB 1; Length 284;  
Best Local Similarity 41.4%; Pred. No. 32;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```

QY      1 KOAEDKVKASREAKKQVEKALEQLLEDKYK 29
       ||||| : | : : : : ||| : 
Db      30 KQAEDRCKQLEEEQQALQKKLGTDEYVE 58

RESULT 27
A23562
tropomyosin 1, fibroblast and epithelial cell - human
C[Species]: Homo sapiens (man)
C[Date]: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 09-Jul-2004
C[Accession]: A23562; J01043
R[MacLeod, A.R.; Houlker, C.; Retnach, F.C.; Smillie, L.B.; Talbot, K.; Modi, G.; Walsh,
Proc. Natl. Acad. Sci. U.S.A. 82, 7835-7839, 1985]
A[Title]: A muscle-type tropomyosin in human fibroblasts: evidence for expression by an a
A[Reference number]: A23562; MUID:86067998; PMID:3865200
A[Accession]: A23562
A[Molecule type]: mRNA
A[Residues]: 1-284 <MAC>
A[Cross-references]: UNIPROT:P07951; GB:M12125; NID:G339951; PIDD:AAA36773.1; PID:G339951
A[Experimental source]: fibroblast
R[Prasad, G.L.; Meisner, S.; Sheer, D.G.; Cooper, H.L.
Biochem. Biophys. Res. Commun. 177, 1068-1075, 1991]
A[Title]: A cDNA encoding a muscle-type tropomyosin cloned from a human epithelial cell 1
A[Reference number]: J01043; MUID:91282744; PMID:2059197
A[Accession]: J01043
A>Status: translation not shown
A[Molecule type]: mRNA
A[Residues]: 1-284 <PRA>
A[Cross-references]: GB:M75165; NID:G339730; PIDD:AAB59509.1; PID:G339731
A[Experimental source]: cell line LS174T
C[Superfamily]: tropomyosin
C[Keywords]: alternative splicing; coiled coil

Query Match          43.1%; Score 59; DB 2; Length 284;
Best Local Similarity 41.4%; Pred. No. 32;
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY      1 KOAEDKVKASREAKKQVEKALEQLLEDKYK 29
       ||||| : | : : : : ||| : 
Db      30 KQAEDRCKQLEEEQQALQKKLGTDEYVE 58

RESULT 28
S00922
tropomyosin beta, skeletal muscle - human
C[Species]: Homo sapiens (man)
C[Date]: 01-Dec-1985 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C[Accession]: S00922
R[Midada, J.S.; Ferraz, C.; Capony, J.P.; Liautard, J.P.
Nucleic Acids Res. 16, 3109, 1988]
A[Title]: Complete nucleotide sequence of the adult skeletal isoform of human skeletal mu
A[Reference number]: S00922; MUID:88217530; PMID:3368322
A[Accession]: S00922
A[Molecule type]: mRNA
A[Residues]: 1-284 <MID>
A[Cross-references]: UNIPROT:P07951; EMBL:X06825; NID:g317248; PIDD:CAA29971.1; PID:g317245
C[Superfamily]: tropomyosin
C[Keywords]: alternative splicing; coiled coil; muscle; skeletal muscle

Query Match          43.1%; Score 59; DB 2; Length 284;
Best Local Similarity 41.4%; Pred. No. 32;
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY      1 KOAEDKVKASREAKKQVEKALEQLLEDKYK 29
       ||||| : | : : : ~ ||| : 
Db      30 KQAEDRCKQLEEEQQALQKKLGTDEYVE 58

RESULT 29
tropomyosin beta 2 - mouse
N[Alternate names]: beta-TM-2
AA4131
```

```

C:\Species: Mus musculus (house mouse)
C:\Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:\Accession: A44131
R:\Wang, Y.C.; Rubenstein, P.A.
J. Biol. Chem. 267, 12004-12010, 1992
A:\Title: Splicing of two alternative exon pairs in beta-tropomyosin pre-mRNA is independent
A:\Reference number: A44131; MUID:92291076; PMID:1601870
A:\Accession: A44131
A:\Molecule type: mRNA
A:\Residues: 1-284 <MAN>
A:\Cross-references: UNIPROT:O61344; GB:M87635; NID:g192156; PIDN:AAA37288.1; PID:g192157
A:\Experimental source: muscle BC3H1 cells
A:\Note: sequence extracted from NCBI backbone (NCBIN:106743, NCBIPI:106744)
C:\Keywords: alternative splicing; coiled coil

Query Match          43.1%; Score 59; DB 2; Length 284;
Best Local Similarity 41.4%; Pred. No. 32;
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY      1 KOAEDKVKASREAKKQVEKALEQLDEDKVK 29
        |||||:::||:::||::||::||:
DB       30 KOAEDRCQLEEEQQALQKKLKTGDEDEVE 58

RESULT 30
S03838
tropomyosin beta, skeletal muscle - mouse
C:\Species: Mus musculus (house mouse)
C:\Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:\Accession: S03838
R:\McInnes, C.; Leader, D.P.
Biochim. Biophys. Acta 951, 117-122, 1988
A:\Title: The tropomyosin mRNAs of mouse striated muscles: molecular cloning of beta-tropo
A:\Reference number: S03838; MUID:89051009; PMID:2461223
A:\Accession: S03838
A:\Molecule type: mRNA
A:\Residues: 1-284 <MCIA>
A:\Cross-references: UNIPROT:P58774; EMBL:X12650; NID:g54856; PIDN:CAA31181.1; PID:g54857
C:\Superfamily: tropomyosin
C:\Keywords: alternative splicing; coiled coil; muscle; skeletal muscle

Query Match          43.1%; Score 59; DB 2; Length 284;
Best Local Similarity 41.4%; Pred. No. 32;
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY      1 KOAEDKVKASREAKKQVEKALEQLDEDKVK 29
        |||||:::||:::||::||::||:
DB       30 KOAEDRCQLEEEQQALQKKLKTGDEDEVE 58

RESULT 31
S23256
tropomyosin beta - mouse
C:\Species: Mus musculus (house mouse)
C:\Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:\Accession: S23256
R;\Wang, Y.C.; Rubenstein, P.A.
J. Biol. Chem. 267, 2728-2736, 1992
A:\Title: Choice of 3' cleavages/polyadenylation site in beta-tropomyosin RNA processing i
A:\Reference number: S23256; MUID:92129366; PMID:1733968
A:\Status: preliminary; translation not shown
A:\Molecule type: mRNA
A:\Residues: 1-284 <MAN>
A:\Cross-references: UNIPROT:P58774; EMBL:X58381; NID:g50189; PIDN:CAA41271.1; PID:g50190
C:\Superfamily: tropomyosin

Query Match          43.1%; Score 59; DB 2; Length 284;
Best Local Similarity 41.4%; Pred. No. 32;
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY      1 KOAEDKVKASREAKKQVEKALEQLDEDKVK 29

```

Db 30 KOAEDRCQLEEEBOALQKCLKTEDEYE 58

## RESULT 32

B25073 tropomyosin beta, skeletal muscle - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C/Accession: B25073  
R/Helfman, D.M.; Cheley, S.; Kuismanen, E.; Finn, L.A.; Yamawaki-Kataoka, Y.  
Mol. Cell. Biol. 6, 3582-3595, 1986  
A/Title: Nonmuscle and muscle tropomyosin isoforms are expressed from a single gene by a  
A/Reference number: A93081; MUID:87089698; PMID:2432392  
A/Accession: B25073  
A/Molecule type: mRNA  
A/Residues: 1-284 <HEL>  
A/Cross-references: UNIPROT:P58775; GB:L00381; GB:M14127; NID:G207494; PIDN:AAA42288.1;  
C/Superfamily: tropomyosin  
C/Keywords: alternative splicing; coiled coil; muscle; skeletal muscle

Query Match 43.1%; Score 59; DB 2; Length 284;

Best Local Similarity 41.4%; Pred. No. 32;

Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KOAEDKYKASREAKKQVEKALEOLEDKYK 29

Db 30 KOAEDRCQLEEEBOALQKCLKTEDEYE 58

## RESULT 33

AH1387 cell wall binding proteins homolog lmo2504 [imported] - listeria monocytogenes (strain E  
C/Species: Listeria monocytogenes  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AH1387  
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Enclian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karet, U.  
Science 294, 849-852, 2001  
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueener, T.; Simoes, N.; Tletter, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,  
A/Title: Comparative genomics of *Listeria species*.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AH1387  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-436 <GLA>  
A/Cross-references: UNIPROT:Q8V4E2; GB:NC\_003210; PIDN:CAD00582.1; PID:G16411992; GSPDB:  
A/Experimental source: strain EGD-e  
C/Genetics:  
A/Gene: lmo2504

Query Match 43.1%; Score 59; DB 2; Length 436;

Best Local Similarity 34.5%; Pred. No. 48;

Matches 10; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KOAEDKYKASREAKKQVEKALEOLEDKYK 29

Db 85 KQEDKVSSENEKUKQKQKEMEKLRNDIR 113

## RESULT 34

B71203 hypothetical protein PH1895 - *Pyrococcus horikoshii*  
C/Species: Pyrococcus horikoshii  
C/Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 09-Jul-2004  
C/Accession: B71203  
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekit  
M.; Ohbuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A/Reference number: A71000; MUID:98344137; PMID:9679194

A/Accession: B71203  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA

A/Residues: 1-217 <KAM>

A/Cross-references: UNIPROT:O59525; GB:AP000007; NID:G3236134; PIDN:BAA31017.1; PID:G325

A/Experimental source: strain OT3

A/Note: this accession replaces an interim accession for a sequence replaced by GenBank

C/Genetics:  
A/Gene: PH1895

Query Match 42.3%; Score 58; DB 2; Length 217;

Best Local Similarity 31.0%; Pred. No. 32;

Matches 9; Conservative 14; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KOAEDKYKASREAKKQVEKALEOLEDKYK 29

Db 112 EKRENEIKQLEKEKLELEIKKIR 140

## RESULT 35

TMCHS1

tropomyosin 1, smooth muscle alpha splice form - chicken

N/Alternate names: gizzard beta-tropomyosin; tropomyosin beta chain, smooth muscle

C/Species: Gallus gallus (chicken)

C/Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 09-Jul-2004

C/Accession: A92462; B31433; I50685; A02982; PC2288

R/Helfman, D.M.; Peramisco, J.R.; Ricci, W.M.; Hughes, S.H.

J. Biol. Chem. 259, 14136-14143, 1984

A/Title: Isolation and sequence of a cDNA clone that contains the entire coding region f

A/Reference number: A92462; MUID:85054861; PMID:6548747

A/Accession: A92462

A/Molecule type: mRNA

A/Residues: 1-284 <HEL>

A/Cross-references: UNIPROT:P19352; UNIPROT:Q9PST0; GB:K02446; NID:G212806; PIDN:AAA4910

R/Sanders, C.; Smillie, L.B.

J. Biol. Chem. 260, 7264-7275, 1985

A/Title: Amino acid sequence of chicken gizzard beta-tropomyosin. Comparison of the chic

A/Reference number: A92537; MUID:85207759; PMID:3997867

A/Accession: A92537

A/Molecule type: protein

A/Residues: 1-284 <SN>

A/Note: the sequence is based in part on homology with rabbit skeletal muscle tropomyosi

R/Libiri, D.; Lemonnier, M.; Meinel, T.; Fiszman, M.Y.

J. Biol. Chem. 264, 2935-2944, 1989

A/Title: A single gene codes for the beta-subunits of smooth and skeletal muscle tropomy

A/Reference number: A31433; MUID:89123396; PMID:2914939

A/Accession: B31433

A/Molecule type: DNA

A/Residues: 1-284 <LIB>

A/Note: the authors translated the codon ATC for residue 19 as Thr, and TTA for residue

R/Sung, L.A.; Lin, J.J.C. Commun. 201, 627-634, 1994

Biochem. Biophys. Res. Commun. 201, 627-634, 1994

A/Title: Erythrocyte tropomodulin binds to the N-terminus of hTMS, a tropomyosin isoform

A/Reference number: PC2283; MUID:94271211; PMID:8002995

A/Contents: annotation

R/Helfman, D.M.; Peramisco, J.R.; Fiddes, J.C.; Thomas, G.P.; Hughes, S.H.

Proc. Natl. Acad. Sci. U.S.A. 80, 31-35, 1983

A/Title: Identification of clones that encode chicken tropomyosin by direct immunologica

A/Reference number: I50685; MUID:83117845; PMID:6185958

A/Accession: I50685

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 151-187 <HE2>

A/Cross-references: EMBL:V00445; NID:G63837; PIDN:CAA23724.1; PID:G929598

C/Genetics:  
A/Gene: TMI

C/Superfamily: tropomyosin  
C/Keywords: actin binding; alternative splicing; blocked amino end; coiled coil

F/I/Modified site: blocked amino end (Met) (probably acetylated) #status experimental

Query Match 42.3%; Score 58; DB 1; Length 284;

Best Local Similarity 41.4%; Pred. No. 41;

Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```
QY      1 KQAEDKVKASREAKKOVEALEQLQLEDKVK 29
          ||||| : ||| : ||| :
DB     30 KQAEDRCQLLEEQQGIGKKLKGTEDEVE 58
```



## RESULT 40

E97357  
hypothetical protein CAC3728 (imported) - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: E97357  
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1c  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: E97357  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-169 <KUR>  
A:Cross-references: UNIPROT:097C08; GB:AE001437; PDB:AAK81648.1; PID:915026835; GSPDB:C  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetic8:  
A:Gene: CAC3728

Query Match 41.6%; Score 57; DB 2; Length 169;  
Best Local Similarity 48.0%; Pred. No. 33;  
Matches 12; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 5 DKVKASREAKKQVEKALQLEDKVK 29  
|||:|||||:|:|:|  
Db 65 DKVEKINEDSKQVKDIYELEDQVK 89

Search completed: June 13, 2005, 20:53:37  
Job time : 33.3721 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:30:50 ; Search time 153.767 Seconds  
(without alignments)  
96.576 Million cell updates/sec

Title: US-10-706-275-2  
Perfect score: 137  
Sequence: 1 KQAEKVKASREAKKQVEKALEQLEDKV 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_tramb1:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.5	54.4	592	2 Q00720	Q00720 streptococc
2	72.5	52.9	454	2 Q840T7	Q840T7 streptococc
3	72.5	52.9	550	2 Q54840	Q54840 streptococc
4	72.5	52.9	587	2 Q55312	Q55312 streptococc
5	72.5	52.6	100	2 Q54639	Q54639 streptococc
6	72.5	52.6	100	2 Q54640	Q54640 streptococc
7	72.5	52.6	100	2 Q54832	Q54832 streptococc
8	72.5	52.6	198	2 Q54832	Q54832 streptococc
9	72.5	52.6	208	2 P95824	P95824 streptococc
10	72.5	52.6	208	2 P95825	P95825 streptococc
11	72.5	52.6	237	2 Q6T1R5	Q6T1R5 streptococc
12	72.5	52.6	251	2 Q6V9Q3	Q6V9Q3 streptococc
13	72.5	52.6	279	2 Q8GL87	Q8GL87 streptococc
14	72.5	52.6	282	2 Q8GLA6	Q8GLA6 streptococc
15	72.5	52.6	303	2 Q8GL98	Q8GL98 streptococc
16	72.5	52.6	307	2 Q8GL94	Q8GL94 streptococc
17	72.5	52.6	314	2 Q8GLB0	Q8GLB0 streptococc
18	72.5	52.6	317	2 Q8GL92	Q8GL92 streptococc
19	72.5	52.6	319	2 Q8GLA2	Q8GLA2 streptococc
20	72.5	52.6	322	2 Q8GLA4	Q8GLA4 streptococc
21	72.5	52.6	326	2 Q8GL91	Q8GL91 streptococc
22	72.5	52.6	340	2 Q8GLA1	Q8GLA1 streptococc
23	72.5	52.6	369	2 Q8GL90	Q8GL90 streptococc
24	72.5	52.6	388	1 PAM_STRPY	PAM_STRPY
25	72.5	52.6	400	2 Q8NZ79	Q8NZ79 streptococc
26	72.5	52.6	408	2 Q54837	Q54837 streptococc
27	72.5	52.6	435	2 Q9AMM3	Q9AMM3 streptococc
28	72.5	52.6	436	2 Q54839	Q54839 streptococc
29	72.5	52.6	441	2 Q55246	Q55246 streptococc
30	72.5	52.6	443	2 Q54703	Q54703 streptococc
31	72.5	52.6	454	2 Q55278	Q55278 streptococc

32	72	52.6	465	2 Q83XW0	Q83XW0 streptococc
33	72	52.6	472	2 Q55279	Q55279 streptococc
34	72	52.6	483	1 M6_STRPY	P08089 streptococc
35	72	52.6	484	2 Q05464	Q05464 streptococc
36	72	52.6	484	2 Q10372	Q10372 streptococc
37	72	52.6	484	2 Q99XV0	Q99XV0 streptococc
38	72	52.6	488	2 Q54830	Q54830 streptococc
39	72	52.6	500	2 Q9RHV2	Q9RHV2 streptococc
40	72	52.6	508	2 Q76MJ2	Q76MJ2 streptococc
41	72	52.6	532	2 Q55098	Q55098 streptococc
42	72	52.6	558	2 Q54718	Q54718 streptococc
43	72	52.6	564	1 M12_STRPY	P19401 streptococc
44	72	52.6	581	2 Q54835	Q54835 streptococc
45	72	52.6	581	2 Q8K5P9	Q8K5P9 streptococc

ALIGNMENTS

RESULT 1					
ID	Q00720	PRELIMINARY;	PRT;	592 AA.	
AC	Q00720;				
DT	01-NOV-1996 (TREMBLrel. 01, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	M protein precursor.				
OC	Streptococcus sp. (lancefield group G).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1320;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=1750;				
RX	MEDLINE=92363566; PubMed=1500178;				
RA	Collins C.M., Kimura A., Bisno A.L.;				
RT	"Group G streptococcal M protein exhibits structural features				
RT	analogous to class I M protein of group A streptococci.";				
RI	Infect. Immun. 60:3689-3696(1992).				
CC	-I- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by				
CC	an amide bond (by similarity).				
DR	EMBL; M95774; AAA26928.1; -				
DR	GO; GO:0009986; C:cell surface; IEA.				
DR	GO; GO:0005618; C:cell wall; IEA.				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	InterPro; IPR011000; Apolp. III like.				
DR	InterPro; IPR005877; Gpos_YsIRK.				
DR	InterPro; IPR001899; Gram_pos_anchor.				
DR	InterPro; IPR003345; M_repeat.				
DR	Pfam; PF00746; Gram_pos_anchor; 1.				
DR	Pfam; PF02370; M_2.				
DR	Pfam; PF04650; YsIRK signal; 1.				
DR	PRINTS; PR00015; GP05ANCHOR.				
DR	TIGRFAMs; TIGR01167; LpxTG_anchor; 1.				
DR	TIGRFAMs; TIGR01168; YsIRK_signal; 1.				
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.				
KW	Cell wall; Peptidoglycan-anchor; Signal.				
FT	SIGNAL 1 41 Potential.				
FT	CHAIN 42 592 M protein.				
SQ	SEQUENCE 592 AA; 67008 MW; 821399D030D5CEB CRC64;				
Query Match 54.4%; Score 74.5; DB 2; Length 592;					
Best Local Similarity 46.3%; Pred. No. 16;					
Matches 19; Conservative 4; Mismatches 5; Indels 13; Gaps 1;					
OY	1 KQAEK-----VKASREAKKQVEKALEQLEDKV 28				
DB	422 KVKEDKQISDASRKGLRDLASREAKKQVEKALEANSKL 462				
RESULT 2					
ID	Q840T7	PRELIMINARY;	PRT;	454 AA.	

```

AC Q8407;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dyll-Smith M.L., Stripakash K.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY263387; AAC92603.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M_1.
FT NON_TER 1
FT TER 454
SQ SEQUENCE 454 AA; 51368 MW; 4DF100DAA6467864 CRC64;

Query Match 52.9%; Score 72.5; DB 2; Length 454;
Best Local Similarity 46.3%; Pred. No. 20;
Matches 19; Conservative 3; Mismatches 6; Indels 13; Gaps 1;

Cy 1 KOAEK-----VKASREAKKQVEKALEQLEDKY 28
Db 381 KVKEDKQISDASRGGLRDLDASREAKKQVEKALEANSKL 421

```

```

RESULT 3
ID Q54840 PRELIMINARY; PRT; 550 AA.
AC Q54840;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M protein.
GN Name=emm55;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Serotype M55;
RL MEDLINE=95371690; PubMed=7643859; DOI=10.1016/0161-5890(95)00022-7;
RA Boyle M.D.P., Weber-Heymann J., Raeder R., Podbielski A.;
RT "Characterization of a gene coding for a type Ito bacterial IgG-binding protein.";
RL Mol. Immunol. 32:669-678(1995).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).
DR EMBL; X72090; CAA50980.1; -
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS08447; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 550 AA; 61736 MW; 32894BFF9805181B CRC64;

```

Query Match 52.9%; Score 72.5; DB 2; Length 550;  
 Best Local Similarity 46.3%; Pred. No. 24;  
 Matches 19; Conservative 3; Mismatches 6; Indels 13; Gaps 1;

```

Cy 1 KOAEK-----VKASREAKKQVEKALEQLEDKY 28
Db 381 KVKEDKQISDASRGGLRDLDASREAKKQVEKALEANSKL 421

```

```

RESULT 4
ID Q55312 PRELIMINARY; PRT; 587 AA.
AC Q55312;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protein V precursor.
GN Name=forV;
OS Streptococcus sp.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1306;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Valente;
RL MEDLINE=93013016; PubMed=1398120; DOI=10.1016/0378-1119(92)90005-A;
RA Smirnov O.Y., Denesyuk A.I., Zakharov M.V., Adamov V.M.,
RA Zayzalov V.P.;
RT "Protein V, a novel type-II IgG receptor from Streptococcus sp.: sequence, homologies and putative Fe-binding site.";
RL Gene 120:27-32(1992).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).
DR EMBL; X62467; CAA44324.1; -
DR PIR; JCI419; JCI419.
DR HSSP; O15813; ID7M.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS08447; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1
FT CHAIN 45
FT POTENTIAL 44
SQ SEQUENCE 587 AA; 64130 MW; D9FA5658AC23FA81 CRC64;

Query Match 52.9%; Score 72.5; DB 2; Length 587;
Best Local Similarity 46.3%; Pred. No. 25;
Matches 19; Conservative 3; Mismatches 6; Indels 13; Gaps 1;

Cy 1 KOAEK-----VKASREAKKQVEKALEQLEDKY 28
Db 417 KVKEDKQISDASRGGLRDLDASREAKKQVEKALEANSKL 457

```

```

RESULT 5
ID Q54639 PRELIMINARY; PRT; 100 AA.
AC Q54639;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M1 protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1314;
RN [1]

```

RP SEQUENCE FROM N.A.  
RA MEDLINE=95172752; PubMed=7868273;  
RX Musser J.M., Kapur V., Szeto J., Pan X., Swanson D.S., Martin D.R.;  
RT "Genetic diversity and relationships among Streptococcus pyogenes  
RT strains expressing serotype M1 protein: recent intercontinental spread  
RT of a subclone causing episodes of invasive disease.";  
RL Infect. Immun. 63:994-1003(1995).  
DR EMBL; U20103; AA85116.1; -.  
DR HSSP; P13276; IEQ1.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 1.  
FT NON\_TER 1 1  
FT NON\_TER 100 100  
SQ SEQUENCE 100 AA; 11243 MW; 9775831FA25DC463 CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 100;  
Best Local Similarity 45.2%; Pred. No. 5.9;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 28 DKVKEKQISDASRQGLRRDLDAASREAKQVEKALEANSKL 69  
RESULT 6  
ID 054640 PRELIMINARY; PRT; 100 AA.  
AC 054640;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE M1 protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95172752; PubMed=7868273;  
RA Musser J.M., Kapur V., Szeto J., Pan X., Swanson D.S., Martin D.R.;  
RT "Genetic diversity and relationships among Streptococcus pyogenes  
RT strains expressing serotype M1 protein: recent intercontinental spread  
RT of a subclone causing episodes of invasive disease.";  
RL Infect. Immun. 63:994-1003(1995).  
DR EMBL; U20104; AA85117.1; -.  
DR HSSP; P13276; IEQ1.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 1.  
FT NON\_TER 1 1  
FT NON\_TER 100 100  
SQ SEQUENCE 100 AA; 11287 MW; 9773331914EDC2D3 CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 100;  
Best Local Similarity 45.2%; Pred. No. 5.9;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 28 DKVKEKQISDASRQGLRRDLDAASREAKQVEKALEANSKL 69  
RESULT 7  
ID 0983A1 PRELIMINARY; PRT; 100 AA.  
AC 0983A1;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 05-JUN-2004 (TRENBLrel. 27, Last annotation update)  
DE M1 protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=95172752; PubMed=7868273;  
RX Musser J.M., Kapur V., Szeto J., Pan X., Swanson D.S., Martin D.R.;  
RT "Genetic diversity and relationships among Streptococcus pyogenes  
RT strains expressing serotype M1 protein: recent intercontinental spread  
RT of a subclone causing episodes of invasive disease.";  
RL Infect. Immun. 63:994-1003(1995).  
DR EMBL; U20102; AA85115.1; -.  
DR EMBL; U20101; AA85114.1; -.  
DR HSSP; P13276; IEQ1.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 1.  
FT NON\_TER 1 1  
FT NON\_TER 100 100  
SQ SEQUENCE 100 AA; 11345 MW; 9773331C0EDC2D3 CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 100;  
Best Local Similarity 45.2%; Pred. No. 5.9;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
Qy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
Db 28 DKVKEKQISDASRQGLRRDLDAASREAKQVEKALEANSKL 69  
RESULT 8  
ID 054832 PRELIMINARY; PRT; 198 AA.  
AC 054832;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm3;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=3-3/317;  
RC MEDLINE=93062420; PubMed=1435517;  
RA Podbielski A., Baird R., Kaufhold A.;  
RT "The group A streptococcal M-type 3 protein gene exhibits a C terminus  
RT typical for class I M proteins.";  
RL Med. Microbiol. Immunol. 181:209-213(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3-3/317;  
RA Podbielski A., Kaufhold A.;  
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
CC an amide bond (by similarity).  
DR EMBL; X66816; CAA47295.1; -.  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF02370; M; 1.  
DR PRINTS; PR00015; GPOSANCHOR.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall, peptidoglycan-anchor.  
FT NON\_TER 1 1  
FT NON\_TER 198 198  
SQ SEQUENCE 198 AA; 21550 MW; A738888D947155D5 CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 198;

Best Local Similarity 45.2%; Pred. No. 11;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQVEKALEQLEEDKV 28  
DB 27 DKVKEKQISDASRQGLRDLDSREAKQVEKALEANSKL 68

## RESULT 9

P95824 PRELIMINARY; PRT; 208 AA.

AC P95824;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=M nontypeable group A;  
RA Brandt E.R., Good M.F.;  
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).

DR EMBL; U65899; AAB4064.1; -  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF02370; M; 1.  
DR PRINTS; PR00015; GPOSANCHOR.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor.

FT NON\_TER 1  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 22695 MW; 013FBBA0A97FF42 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 208;  
Best Local Similarity 45.2%; Pred. No. 11;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQVEKALEQLEEDKV 28  
DB 37 DKVKEKQISDASRQGLRDLDSREAKQVEKALEANSKL 78

## RESULT 10

P95826 PRELIMINARY; PRT; 208 AA.

AC P95826;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=RHDI52-;  
RA Brandt E.R., Good M.F.;  
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).  
DR EMBL; U66005; AAB40642.1; -  
DR GO; GO:0009986; C:cell surface; IEA.

DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF02370; M; 1.  
DR PRINTS; PR00015; GPOSANCHOR.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor.

FT NON\_TER 1  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 22695 MW; 79972A987324729B CRC64;

Query Match 52.6%; Score 72; DB 2; Length 208;  
Best Local Similarity 45.2%; Pred. No. 11;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQVEKALEQLEEDKV 28  
DB 37 DKVKEKQISDASRQGLRDLDSREAKQVEKALEANSKL 78

## RESULT 11

O6TLR5 PRELIMINARY; PRT; 237 AA.

AC O6TLR5;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;

OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=NS241;  
RX MEDLINE=22894607; PubMed=14532198;  
RA Moses A.E., Hidalgo-Grass C., Dan-Goor M., Jaffe J., Shetzigovsky I., Ravine M., Korenman Z., Cohen-Poradosu R., Nir-Paz R.;  
RT "emm typing of M nontypeable invasive group A streptococcal isolates in Israel.";  
RL J. Clin. Microbiol. 41:4655-4659(2003).  
DR EMBL; AY394538; AAC94530.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 1.

FT NON\_TER 1  
FT NON\_TER 237  
SQ SEQUENCE 237 AA; 27027 MW; 47CF315DD4B5F2 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 237;  
Best Local Similarity 42.5%; Pred. No. 13;  
Matches 17; Conservative 6; Mismatches 5; Indels 12; Gaps 1;

QY 1 KOAEDKVK-----ASREAKQVEKALEQLEEDKV 28  
DB 172 KBQNKISBASRQGLRDLDSREAKQVEKALEANSKL 211

## RESULT 12

O6V9Q3 PRELIMINARY; PRT; 251 AA.

AC O6V9Q3;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;

OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```

OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J58;
RA Moses A.E., Hidalgo-Grass C., Dan-Goor M., Jaffe J., Shatzgovsky I.,
RL Rabin M., Korenman Z., Cohen-Poradosu R., Nir-Paz R.;
DR Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY346386; AAQ73206.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1
FT SEQUENCE 251 AA; 28938 MW; 2A6602AAA637D11 CRC64;
SQ

Query Match 52.6%; Score 72; DB 2; Length 251;
Best Local Similarity 45.2%; Pred. No. 13;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQYKALEQLEDKV 28
Db 199 DKVKEKQISDASRQGLRRDLDA$REAKQYKALEANSKL 240

RESULT 13
Q8GL87 PRELIMINARY; PRT; 279 AA.
AC O8GL87;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dvall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,
RL McMillan D.;
DR Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY139420; AAN64693.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 2.
FT NON_TER 1
FT SEQUENCE 279 AA; 31224 MW; 16A600455B5C3A0 CRC64;
SQ

Query Match 52.6%; Score 72; DB 2; Length 279;
Best Local Similarity 45.2%; Pred. No. 15;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQYKALEQLEDKV 28
Db 202 DKVKEKQISDASRQGLRRDLDA$REAKQYKALEANSKL 243

RESULT 14
Q8GLA6 PRELIMINARY; PRT; 282 AA.
AC O8GLA6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Dvall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,
RL McMillan D.;
DR Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY139401; AAN64674.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 1.
FT NON_TER 1
FT SEQUENCE 282 AA; 32157 MW; 5C799B0AA4323541 CRC64;
SQ

Query Match 52.6%; Score 72; DB 2; Length 282;
Best Local Similarity 45.2%; Pred. No. 15;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQYKALEQLEDKV 28
Db 204 DKVKEKQISDASRQGLRRDLDA$REAKQYKALEANSKL 245

RESULT 15
Q8GL98 PRELIMINARY; PRT; 303 AA.
AC O8GL98;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dvall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,
RL McMillan D.;
DR Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY139409; AAN64682.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M; 2.
FT NON_TER 1
FT SEQUENCE 303 AA; 34562 MW; F76F37540E16CD1B CRC64;
SQ

Query Match 52.6%; Score 72; DB 2; Length 303;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 16; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

Qy 1 KOAED-----KVKA$REAKQYKALEQLEDKV 28
Db 230 KOVSDASRQGLRRDLDA$REAKQYKALEANSKL 265

RESULT 16
Q8GL84 PRELIMINARY; PRT; 307 AA.
AC O8GL84;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dvall-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,

```

```
RA McMillan D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY139423; AAN64696.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF02370; M; 2.
FT NON_TER 1
SQ SEQUENCE 307 AA; 34955 MW; 226822938E6E0E CRC64;

Query Match
Best Local Similarity 52.6%; Score 72; DB 2; Length 307;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 229 DKVKEKQISDASRQGLRRDLDASREAKKQVEKALEFANSKL 270

RESULT 17
O8GLB0 PRELIMINARY; PRT; 314 AA.
AC O8GLB0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dyall-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY139397; AAN64670.1; -.
DR HSSP: P04268; 1IC2.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF02370; M; 2.
FT NON_TER 1
FT NON_TER 314
SQ SEQUENCE 314 AA; 35613 MW; 66173BBEB74C9EFC CRC64;

Query Match
Best Local Similarity 52.6%; Score 72; DB 2; Length 314;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 237 DKVKEKQISDASRQGLRRDLDASREAKKQVEKALEFANSKL 278

RESULT 18
O8GL92 PRELIMINARY; PRT; 317 AA.
AC O8GL92;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dyall-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR McMillan D.;
```

```
DR EMBL: AY139415; AAN64688.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF02370; M; 2.
FT NON_TER 1
FT NON_TER 317
SQ SEQUENCE 317 AA; 36056 MW; F2026105F02D4888 CRC64;

Query Match
Best Local Similarity 52.6%; Score 72; DB 2; Length 317;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 240 DKVKEKQISDASRQGLRRDLDASREAKKQVEKALEFANSKL 281

RESULT 19
O8GLA2 PRELIMINARY; PRT; 319 AA.
AC O8GLA2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dyall-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY139405; AAN64678.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF02370; M; 2.
FT NON_TER 1
FT NON_TER 319
SQ SEQUENCE 319 AA; 36283 MW; 1797D2026105F02D CRC64;

Query Match
Best Local Similarity 52.6%; Score 72; DB 2; Length 319;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 240 DKVKEKQISDASRQGLRRDLDASREAKKQVEKALEFANSKL 281

RESULT 20
O8GLA4 PRELIMINARY; PRT; 322 AA.
AC O8GLA4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Dyall-Smith M.L., Krejany S., Sriprakash K.S., Delvecchio A.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY139403; AAN64676.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR003345; M_repeat.
```



DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
SQ SEQUENCE 322 AA; 36990 MW; F179D8DF126CED92 CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 322;  
Best Local Similarity 45.2%; Pred. No. 17;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
OY 5 DKVK-----ASREAKQVEKALEQLEDKY 28  
DB 243 DKVKEKQISDASRQGLRRDLASREAKQVEKALEANSKL 284

RESULT 21  
ID O8GL91 PRELIMINARY; PRT; 326 AA.  
AC O8GL91;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A. Krejany S., Striprakash K.S., Delvecchio A.,  
RA Dyal1-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139416; AAN64689.1; -.  
DR HSSP; P02647; IAV1.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
SQ SEQUENCE 326 AA; 37290 MW; 7186F85FBE9970C7 CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 326;  
Best Local Similarity 45.2%; Pred. No. 17;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
OY 5 DKVK-----ASREAKQVEKALEQLEDKY 28  
DB 249 DKVKEKQISDASRQGLRRDLASREAKQVEKALEANSKL 290

RESULT 22  
ID O8GLA1 PRELIMINARY; PRT; 340 AA.  
AC O8GLA1;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A. Krejany S., Striprakash K.S., Delvecchio A.,  
RA Dyal1-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139406; AAN64679.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1

FT NON\_TER 340  
SQ SEQUENCE 340 AA; 38495 MW; 2222F0DE240268CD CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 340;  
Best Local Similarity 45.2%; Pred. No. 17;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
OY 5 DKVK-----ASREAKQVEKALEQLEDKY 28  
DB 261 DKVKEKQISDASRQGLRRDLASREAKQVEKALEANSKL 302

RESULT 23  
ID O8GL90 PRELIMINARY; PRT; 369 AA.  
AC O8GL90;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE M protein (Fragment).  
GN Name=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A. Krejany S., Striprakash K.S., Delvecchio A.,  
RA Dyal1-Smith M.L., Krejany S., Striprakash K.S., Delvecchio A.,  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY139417; AAN64690.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF02370; M; 2.  
FT NON\_TER 1  
SQ SEQUENCE 369 AA; 41998 MW; EA81172C872012ED CRC64;  
Query Match 52.6%; Score 72; DB 2; Length 369;  
Best Local Similarity 45.2%; Pred. No. 19;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
OY 5 DKVK-----ASREAKQVEKALEQLEDKY 28  
DB 292 DKVKEKQISDASRQGLRRDLASREAKQVEKALEANSKL 333

RESULT 24  
ID PAM\_STRPY STANDARD; PRT; 388 AA.  
AC P49054;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Plasmidogen-binding group A streptococcal M-like protein PAM precursor (Fragment).  
GN Name=pam; Synonyms=emm;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A. AND SEQUENCE OF 32-41.  
RC STRAIN=AP53 / Serotype M53;  
RX MEDLINE=9406405; PubMed=8244975;  
RA Berge A., Sjoebing U.;  
RT "PAM, a novel plasmidogen-binding protein from Streptococcus pyogenes";  
RL J. Biol. Chem. 268:25417-25424 (1993).  
RN [2]  
RP CHARACTERIZATION OF PLASMINOGEN BINDING AND MUTAGENESIS.  
RC STRAIN=AP53 / Serotype M53;  
RX PubMed=8748039;

```

RA Carlsson Wistedt A., Ringdahl U., Mueller-Esterl W., Sjoebirg U.;
RT "Identification of a plasminogen-binding motif in PAM, a bacterial
RT surface protein.";
RL Mol. Microbiol. 18:569-578(1995).
CC
CC -1- FUNCTION: Binds to human plasminogen (and plasmin) via its kringle
CC repeats. Also binds to albumin, immunoglobulin G and fibrinogen.
CC Could provide the bacteria with a mechanism for invasion, as
CC streptococcal-bound plasmin could permit tissue penetration.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- MISCELLANEOUS: PAM has more than one binding site for plasminogen;
CC it is thought that each of the a-repeats can bind one plasminogen
CC molecule.
CC -1- SIMILARITY: Belongs to the M protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22219; CAA80222.1; -.
DR PIR; A49545; A49545.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF02370; M_2.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; PARTIAL.
KW Cell wall, Direct protein sequencing; Peptidoglycan-anchor; Repeat;
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 29 Potential.
FT CHAIN 30 384 Plasminogen-binding group A streptococcal
FT PROPEP 385 >388 M-like protein PAM.
FT DOMAIN 85 113 Removed by sortase (Potential).
FT REPEAT 91 116 Able to bind plasminogen.
FT REPEAT 104 103 2 X approximate tandem repeats, type a.
FT REPEAT 104 103 A-1.
FT REPEAT 147 161 A-2.
FT REPEAT 147 161 2 X tandem repeats, type b.
FT REPEAT 147 153 B-1.
FT REPEAT 154 161 B-2.
FT DOMAIN 163 278 3 X tandem repeats, type c.
FT REPEAT 163 204 C-1.
FT REPEAT 205 246 C-2.
FT REPEAT 247 280 C-3 (Incomplete).
FT DOMAIN 344 380 Gly/Pro-rich.
FT SITE 381 385 LPXTG sorting signal (Potential).
FT MOD_RES 384 384 Pentaglycyl murein peptidoglycan amidated
threonine (Potential).
FT MUTAGEN 82 82 K->A: No change in plasminogen binding.
FT MUTAGEN 98 98 K->A: 50-fold decrease in plasminogen
binding.
FT MUTAGEN 111 111 K->A: 2-fold decrease in plasminogen
binding.
FT NON_TER 388 388
FT SEQUENCE 388 AA; 43629 MW; EEECAFD962CCDB12 CRC64;
SQ
Query Match 52.6%; Score 72; DB 1; Length 388;
Best Local Similarity 45.2%; Pred. No. 20;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

```

```

ID Q8N279 PRELIMINARY; PRT; 400 AA.
AC Q8N279;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M18 protein.
GN Name=emm18; Ordered locusName=spyM18_2076;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barlian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Studevant D.E., Ricklets S.M., Porcella S.P.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Yeasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (by similarity).
CC -----
DR EMBL; AE010111; ALA98546.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall, Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 400 AA; 44645 MW; D73F2A1A8BC781F CRC64;

```

Query Match 52.6%; Score 72; DB 2; Length 400;  
Best Local Similarity 45.2%; Pred. No. 20;  
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

```

Qy 5 DKVK-----ASREAKQVKEKALBQLEDKV 28
Db 230 DKVKEKQISDASRQGLRDLDSREAKQVKEKALBANSKL 271

RESULT 26
ID Q54837 PRELIMINARY; PRT; 408 AA.
AC Q54837;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M protein type 41.
GN Name=emm41;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C101/103/2;
RX MEDLINE=93204905; PubMed=8455563;
RA Podbielski A.;
RT "Three different types of organization of the vir regulon in group A
RT streptococci.";
RL Mol. Gen. Genet. 237:287-300(1993).

```

CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (by similarity).  
 DR EMBL; X58178; CAA41167.1; -.  
 DR PIR; S30283; S30283.  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR011000; Apolp\_III\_like.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M; 2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LpxTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall, peptidoglycan-anchor.  
 SQ SEQUENCE 408 AA; 45630 MW; 5F449A7645BB096C CRC64;

Query Match 52.6%; Score 72; DB 2; Length 408;  
 Best Local Similarity 50.0%; Pred. No. 21;  
 Matches 18; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

Oy 1 KOAED-----KYKASREAKQVEKALEQLEDKV 28  
 Db 243 KOVSASRQGLRDLDAAREAKQVEKALEANSKL 278

## RESULT 27

O9AMW3 PRELIMINARY; PRT; 435 AA.

ID O9AMW3  
 AC O9AMW3  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE M ST4547 protein.  
 GN Name=emmST4547;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 KW NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Behaghi M., Ali M., Yusoff K., Jamal F.;  
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (by similarity).  
 DR EMBL; AF34061; AAK11612.1; -.  
 DR HSBP; P02647; IAV1.  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M; 2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LpxTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall, peptidoglycan-anchor.  
 SQ SEQUENCE 435 AA; 49022 MW; A6C48F7137AE046 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 435;  
 Best Local Similarity 45.2%; Pred. No. 22;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
 Oy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
 Db 5 DKVK-----ASREAKQVEKALEQLEDKV 28

Db 271 DKVEKQISDASRQGLRDLDAAREAKQVEKALEANSKL 312

## RESULT 28

ID O54839 PRELIMINARY; PRT; 436 AA.

AC O54839  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE M protein type 52.  
 GN Name=emm52;  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 KW NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A871/14/3;  
 RX MEDLINE=93204905; PubMed=8455563;  
 RA Podbielski A.;  
 RT "three different types of organization of the vir regulon in group A  
 streptococci.";  
 RL Mol. Gen. Genet. 237:287-300(1993).  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (by similarity).  
 DR EMBL; X58179; CAA41168.1; -.  
 DR PIR; S30284; S30284.  
 DR GO; GO:0009986; C:cell surface; IEA.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR005877; Gpos\_Ysirk.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M; 2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01167; LpxTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall, peptidoglycan-anchor.  
 SQ SEQUENCE 436 AA; 48951 MW; 4C5720F98F2DAE89 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 436;  
 Best Local Similarity 45.2%; Pred. No. 22;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Oy 5 DKVK-----ASREAKQVEKALEQLEDKV 28  
 Db 265 DKVEKQISDASRQGLRDLDAAREAKQVEKALEANSKL 306

## RESULT 29

ID O55246 PRELIMINARY; PRT; 441 AA.

AC O55246  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE M protein.  
 GN Name=emm;  
 OS Streptococcus sp.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 KW NCBI\_TaxID=1306;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Group G;  
 RX MEDLINE=95229925; PubMed=7714192;  
 RA Schmitzler N., Podbielski A., Baumgarten G., Mignon M., Kaufhold A.;  
 RT "M or M-like protein gene polymorphisms in group G Streptococci.";  
 RL J. Clin. Microbiol. 33:356-363(1995).

```
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL: X60098; CAA42694.1; -.
DR GO: GO:0009986; C:cell surface; IEA.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR011000; Apolp_III like.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02370; M; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR PRINTS: PR00015; GPOSANCHOR.
DR TIGRFAMs: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall, peptidoglycan-anchor.
SQ SEQUENCE 441 AA; 49868 MW; 33DPF991E365D9455 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 441;
Best Local Similarity 45.2%; Pred. No. 22;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLEDRV 28
Db 270 DKVKEKQISDASRKGLRDLDAAREAKQVEKALEANSKLT 311

RESULT 30
ID 054703 PRELIMINARY; PRT; 443 AA.
AC 054703;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Emml18.1 protein.
GN Name=emml18.1;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type 18;
RX MEDLINE=96201553; PubMed=8613352;
RA Dale J.B., Washburn R.G., Marques M.B., Messels M.R.;
RT "Hyaluronate capsule and surface M protein in resistance to
RT opsonization of group A streptococci.";
RL Infect. Immun. 64:1495-1501(1996).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL: U29585; AAB03086.1; -.
DR GO: GO:0009986; C:cell surface; IEA.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR011000; Apolp_III like.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02370; M; 3.
DR Pfam: PF04650; Ysirk_signal; 1.
DR PRINTS: PR00015; GPOSANCHOR.
DR TIGRFAMs: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall, peptidoglycan-anchor.
SQ SEQUENCE 443 AA; 49365 MW; 239B0E7F0ADB5EAB CRC64;
```

```
Qy 5 DKVK-----ASREAKQVEKALEQLEDRV 28
Db 272 DKVKEKQISDASRKGLRDLDAAREAKQVEKALEANSKLT 313

RESULT 31
ID 055278 PRELIMINARY; PRT; 454 AA.
AC 055278;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MLC36 precursor (Fragment).
OS Streptococcus sp. 'group C'.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=33972;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C36;
RA Ben Nasr A., Wistedt A., Ringdahl U., Sjobring U.;
RT "Streptokinase activates plasminogen bound to human group C and group
RT G streptococci through M-like proteins.";
RL Eur. J. Biochem. 0:0-0(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C36;
RX MEDLINE=94291620; PubMed=8020465;
RA Ben Nasr A., Wistedt A., Ringdahl U., Sjobring U.;
RT "Streptokinase activates plasminogen bound to human group C and group
RT G streptococci through M-like proteins.";
RL Eur. J. Biochem. 222:267-276(1994).
DR EMBL: J32677; CAA83588.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF02370; M; 4.
DR Pfam: PF04650; Ysirk_signal; 1.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 29 Potential.
FT CHAIN 30 >454 MLC36.
FT NON_TER 454 454
SQ SEQUENCE 454 AA; 51416 MW; FA7D34562548282F CRC64;

Query Match 52.6%; Score 72; DB 2; Length 454;
Best Local Similarity 45.2%; Pred. No. 23;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Qy 5 DKVK-----ASREAKQVEKALEQLEDRV 28
Db 310 DKVKEKQISDASRKGLRDLDAAREAKQVEKALEANSKLT 351

RESULT 32
ID 083XW0 PRELIMINARY; PRT; 465 AA.
AC 083XW0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M protein (Fragment).
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22894661; PubMed=14532252;
RA Lau S.K., Woo P.C., Yim T.C., To A.P., Yuen K.Y.;
RT "Molecular characterization of a strain of group A streptococcus
```

RT Isolated from a patient with a proca abscesses";  
 RL J. Clin. Microbiol. 41:4888-4891(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lau S.K.P., Woo P.C.Y., Yim T.C., To A.P.C., Yuen K.Y.;  
 RL Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 an amide bond (By similarity).  
 DR EMBL: AY273148; AAP32473.1; -.  
 DR HSBP; P04268; 1IC2.  
 DR GO: GO:0009986; C:cell surface; IEA.  
 DR GO: GO:0005618; C:cell wall; IEA.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR InterPro: IPR011000; APOLP\_III like.  
 DR InterPro: IPR005877; Gpos\_Ysirk.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR003345; M\_repeat.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M; 3.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 DR Cell wall; peptidoglycan-anchor.  
 RN NON\_TER  
 FT 1  
 SQ SEQUENCE 465 AA; 52098 MW; 5971795A94C6B78 CRC64;  
 Query Match 52.6%; Score 72; DB 2; Length 465;  
 Best Local Similarity 45.2%; Pred. No. 23;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
 Oy 5 DKVK-----ASREAKQYKALEQLEDKY 28  
 |||||  
 Db 305 DKVKEKQISDPSRKGRLRDLDASREAKQYKALEANSKL 346  
 |||||  
 RESULT 33  
 Q55279 PRELIMINARY; PRT; 472 AA.  
 AC Q55279;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE MLG72 precursor.  
 OS Streptococcus sp. (Lancefield group G).  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus;  
 CC NCBI\_TaxID=1320;  
 OK [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Q72;  
 RC MEDLINE=94291620; PubMed=8020466;  
 RX Ben Naer A., Wistred A., Ringdahl U., Stobring U.;  
 RT "Streptokinase activase plasminogen bound to human group C and group  
 RL G streptococci through M-like proteins";  
 DL Eur. J. Biochem. 222:267-276(1994).  
 DR EMBL; Z32678; CAA83589.1; -.  
 DR HSBP; O9JNH0; 1G04.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR InterPro: IPR005877; Gpos\_Ysirk.  
 DR InterPro: IPR003345; M\_repeat.  
 DR Pfam; PF02370; M; 4.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR Signal.  
 FT 1  
 FT CHAIN 1 29 Potential.  
 FT PROPEP 30 472 Mature MLG72.  
 FT DOMAIN 472 AA; 53968 MW; 8DE0886B2F45FFC8 CRC64;  
 SQ SEQUENCE 472 AA; 53968 MW; 8DE0886B2F45FFC8 CRC64;  
 Query Match 52.6%; Score 72; DB 2; Length 472;  
 Best Local Similarity 45.2%; Pred. No. 23;  
 Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;  
 Oy 5 DKVK-----ASREAKQYKALEQLEDKY 28

Db 328 DKVKEKQISDPSRKGRLRDLDASREAKQYKALEANSKL 369  
 |||||  
 RESULT 34  
 ID M6 STREPY STANDARD; PRT; 483 AA.  
 AC P06089;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE M protein, serotype 6 precursor.  
 GN Name=emm6;  
 OS Streptococcus pyogenes.  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus;  
 CC NCBI\_TaxID=1314;  
 OK [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=8611835; PubMed=3511046;  
 RX Hollingshead S.K., Fischetti V.F., Scott J.R.;  
 RT "Complete nucleotide sequence of type 6 M protein of the group A  
 RL Streptococcus. Repetitive structure and membrane anchor";  
 DL J. Biol. Chem. 261:1677-1686(1986).  
 RN [2]  
 RP SEQUENCE OF 43-122 FROM N.A.  
 RX MEDLINE=85166224; PubMed=3885219;  
 RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;  
 RT "Relationship of M protein genes in group A streptococci";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).  
 CC -1- FUNCTION: This protein is one of the different antigenic serotypes  
 of protein M. Protein M is closely associated with virulence of  
 the bacterium and can render the organism resistant to  
 phagocytosis.  
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 an amide bond (potential).  
 CC -1- SIMILARITY: Belongs to the M protein family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M11338; AAA26920.1; -.  
 DR PIR; A26297; A26297.  
 DR InterPro: IPR011000; APOLP\_III like.  
 DR InterPro: IPR005877; Gpos\_Ysirk.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR003345; M\_repeat.  
 DR InterPro: IPR003611; Inton\_nuc\_2.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF02370; M; 2.  
 DR Pfam; PF04650; Ysirk\_signal; 1.  
 DR PRINTS; PR00015; GPOSANCHOR.  
 DR SMART; SM00496; IENR2; 4.  
 DR TIGRFAMs; TIGR01167; LpxTG\_anchor; 1.  
 DR TIGRFAMs; TIGR01168; Ysirk\_signal; 1.  
 DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
 KW Antigen; Cell wall; Coiled coil; Peptidoglycan-anchor; Phagocytosis;  
 KW Repeat; Signal; Virulence.  
 FT 1  
 FT SIGNAL 1 42  
 FT CHAIN 43 452 M protein, serotype 6.  
 FT PROPEP 453 483 Removed by sortase (potential).  
 FT DOMAIN 69 138 10 X 7 AA approximate tandem repeats of  
 [KMR]-L-[TQ]-[TDA]-[ENQ]-N-[NDK].  
 FT REPEAT 69 75 1-1.  
 FT REPEAT 76 82 2-1.  
 FT REPEAT 83 89 3-1.  
 FT REPEAT 90 96 4-1.  
 FT REPEAT 97 103 5-1.

```
FT REPEAT 104 110 6-1.
FT REPEAT 111 117 7-1.
FT REPEAT 118 124 8-1.
FT REPEAT 125 131 9-1 (approximate).
FT REPEAT 132 138 10-1.
FT REPEAT 138 144 4.5 X 25 AA tandem repeats of E-(NS)-K-E-
  (TA)-I-G-T-L-K-K-(TI)-L-D-E-T-V-K-D-K-I-
  A-[KR]-E-Q.
FT REPEAT 157 181 1-2.
FT REPEAT 182 206 2-2.
FT REPEAT 207 231 3-2.
FT REPEAT 232 256 4-2.
FT REPEAT 257 281 5-2 (approximate).
FT DOMAIN 279 347 Two directly repeated 27 amino acid
  blocks separated by 15 amino acids.
FT DOMAIN 348 411 Hydrophilic.
FT DOMAIN 412 448 Gly/Pro-rich; cell wall-spanning.
FT SITE 449 453 LpXTG sorting signal (Potential).
FT MOD_RES 452 452 Pentaglycyl murein peptidoglycan amidated
  threonine (Potential).
SQ SEQUENCE 483 AA; 53472 MW; 68f87f28b53a448 CRC64;

Query Match 52.6%; Score 72; DB 1; Length 483;
Best Local Similarity 45.2%; Pred. No. 24;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 312 DKVKEKQISDASRQGLRDLDA$REAKKQVEKALEANSKL 353

RESULT 35
ID 005464 PRELIMINARY; PRT; 484 AA.
AC 005464.
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DE M1.1 protein precursor.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS190;
RX MEDLINE=93360826; PubMed=8355619;
  Harbaugh M.P., Podbielski A., Huegl S., Cleary P.P.;
  "Nucleotide substitutions and small-scale insertion produce size and
  antigenic variation in group A streptococcal M1 protein."
RT Mol. Microbiol. 8:981-991(1993).
RL M1.1 SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
  an amide bond (By similarity).
CC EMBL; Z21845; CAA79893.1; -.
DR PIR; S34978; S34978.
DR PIR; S46489; S46489.
DR HSSP; P04268; IIC2.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSTRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_3.
DR Pfam; PF04650; YSTRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSTRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
  Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 41 Potential.
FT CHAIN 42 484 M1.1 protein.
```

```
SQ SEQUENCE 484 AA; 54341 MW; 2B310CC71F8100CF CRC64;

Query Match 52.6%; Score 72; DB 2; Length 484;
Best Local Similarity 45.2%; Pred. No. 24;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 313 DKVKEKQISDASRQGLRDLDA$REAKKQVEKALEANSKL 354

RESULT 36
ID 010372 PRELIMINARY; PRT; 484 AA.
AC 010372.
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DE M protein type 1.
GN Name=emm1;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX Harbaugh M.P., Podbielski A., Huegl S., Cleary P.P.;
  "Nucleotide substitutions lead to antigenic variation in the V-
  terminal non-repeat domains of group A streptococcal M1 protein."
RT Mol. Microbiol. 7:0-0(1993).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93360826; PubMed=8355619;
  Harbaugh M.P., Podbielski A., Huegl S., Cleary P.P.;
  "Nucleotide substitutions and small-scale insertion produce size and
  antigenic variation in group A streptococcal M1 protein."
RT Mol. Microbiol. 8:981-991(1993).
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RX Podbielski A.;
  Submitted (SEP-1991) to the EMBL/GenBank/DBJ databases.
RL M1.1 SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
  an amide bond (By similarity).
CC EMBL; X62131; CAA44062.1; -.
DR PIR; S35401; S35401.
DR PIR; S46489; S46489.
DR HSSP; P04268; IIC2.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSTRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_3.
DR Pfam; PF04650; YSTRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSTRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
  Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 484 AA; 54234 MW; 4179A7BBA5FAED9 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 484;
Best Local Similarity 45.2%; Pred. No. 24;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

Cy 5 DKVK-----ASREAKKQVEKALEQLEDKV 28
Db 313 DKVKEKQISDASRQGLRDLDA$REAKKQVEKALEANSKL 354

RESULT 37
```

```

Q99XV0
ID Q99XV0 PRELIMINARY; PRT; 484 AA.
AC Q99XV0;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M protein type 1.
GN Name=emm1; OrderedLocuNames=SPY2018;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1314;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RA MEDLINE=2192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Ferretti J.J., Moshen W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Serate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (by similarity).
DR EMBL; AB006624; AKK34694.1; -.
DR PIR; S46489; S46489.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR05877; Gpos_Ysirk.
DR InterPro; IPR01899; Gram_pos_anchor.
DR InterPro; IPR03345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_3.
DR Pfam; PF04650; Ysirk_signal; 1.
DR PRINTS; PRO0015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
KM SEQUENCE 484 AA; 54220 MW; 99DD63DDE3799F3C CRC64;
SQ
Query Match 52.6%; Score 72; DB 2; Length 484;
Best Local Similarity 45.2%; Pred. No. 24;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQYKALEQLEEDV 28
Db 313 DKVKEKQISDASRQGLRDLDAEREAKQYKALEEANSKL 354

RESULT 38
ID Q54830 PRELIMINARY; PRT; 488 AA.
AC Q54830;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M 3 protein precursor (fragment).
GN Name=emm3;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1314;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C203;
RA MEDLINE=94267304; PubMed=8207299;
RA Katsukawa C.;
RT "Cloning and nucleotide sequence of the type 3 M protein gene (emm3)
RT consisting of an N-terminal variable portion and C-terminal conserved
RT C-repeat regions: Relation to other genes of Streptococcus pyogenes.";
```

```

RL Kansenshogaku Zasshi 68:698-705(1994).
DR EMBL; D14415; BAA0311.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR05877; Gpos_Ysirk.
DR InterPro; IPR03345; M_repeat.
DR Pfam; PF02370; M_2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KM Signal.
FT NON_TER 1 1
FT SIGNAL <11 32 Potential.
FT CHAIN 33 481 M 3 protein.
FT NON_TER 488 488
SQ SEQUENCE 488 AA; 55208 MW; 04F49300BAF21363 CRC64;

Query Match 52.6%; Score 72; DB 2; Length 488;
Best Local Similarity 45.2%; Pred. No. 24;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;

QY 5 DKVK-----ASREAKQYKALEQLEEDV 28
Db 401 DKVKEKQISDASRQGLRDLDAEREAKQYKALEEANSKL 442

RESULT 39
ID Q9RHV2 PRELIMINARY; PRT; 500 AA.
AC Q9RHV2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE M2.
GN Name=emm23;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1314;
RN [1]
SEQUENCE FROM N.A.
RM MEDLINE=20187076; PubMed=10724481; DOI=10.1016/S0923-2508(00)00130-3;
RX Hong K.;
RT "Characterization of group A streptococcal strains 5v and 5u:
RT determination of emm gene typing and presence of small vir region.";
RL Res. Microbiol. 151:25-36(2000).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (by similarity).
DR EMBL; AB016537; BAA36703.2; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR05877; Gpos_Ysirk.
DR InterPro; IPR01899; Gram_pos_anchor.
DR InterPro; IPR03345; M_repeat.
DR InterPro; IPR00533; Tropomyosin.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M_2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR PRINTS; PRO0015; GPOSANCHOR.
DR PRINTS; PRO0194; TROPOMYOSIN.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Peptidoglycan-anchor.
KM SEQUENCE 500 AA; 55968 MW; 463CA544B3EDE3F0 CRC64;
SQ
Query Match 52.6%; Score 72; DB 2; Length 500;
Best Local Similarity 45.2%; Pred. No. 25;
Matches 19; Conservative 2; Mismatches 3; Indels 18; Gaps 1;
```

